Calculus 3

Course Notes for MATH 237

Edition 7.0

J. Wainwright, J. West, D. Wolczuk

Contents

Preface

Content:

This text covers the material of a traditional first course in multivariable calculus, apart from vector integral calculus, which is contained in the course Calculus 4 (AMath 231).

Prerequisites:

- A good knowledge of the fundamentals of one-variable calculus (limits, differentiation, the Chain Rule, the linear approximation, Taylor polynomials, curve sketching, the Riemann integral ...).
- A good knowledge of the fundamentals of linear algebra (vector algebra, matrix algebra, linear mappings, and determinants).

When studying multivariable calculus one begins to see how the concepts of linear algebra begin to interact with those of calculus.

Why is Calculus 3 a core course?

Multivariable calculus is one of the basic tools in the mathematical sciences. The material in this course is used in a variety of 3rd and 4th year courses in all departments of the Faculty of Mathematics. Examples of subject areas and related courses for which Calculus 3 is a prerequisite are:

- ordinary and partial differential equations (AMATH 351, 353)
- introduction to optimization (CO 255)
- non-linear programming (CO 367)
- introduction to computational mathematics (CS 371)
- mathematical statistics (STAT 330)
- real and complex analysis (PMATH 331, 332)

Viewpoint:

In writing this text, we have emphasized three aspects of multivariable calculus:

- the geometrical interpretation
- analytical computational skills
- the formal theoretical aspects (definitions, theorems and proofs)

Applications are mentioned as motivation, but are not discussed in depth.

We have given formal definitions of all concepts and have given precise statements of all theorems. In the first ten chapters we have given detailed proofs of most of the important theorems (the Differentiability Theorem, the Chain Rule, Taylor's theorem, and the Second Derivative Test). There are fewer formal proofs in the final four chapters, with almost none in Chapters 11 and 14, although the theorems are justified heuristically. We have taken care to make a clear distinction between a formal proof and a heuristic argument.

Most of the concepts are discussed primarily for functions of two variables, but the case for more than two variables is usually discussed at the end of a section, under the heading "generalization".

To the Student Reader

This text is written for students who are willing to work hard in order to obtain a good understanding of multivariable calculus, so as to be able to apply the concepts and methods elsewhere. In order to be successful in this course it is essential to know single variable calculus well. So keep your notes from Math 137/138 or your first year calculus text handy for reference.

This text is intended to be studied with a pencil and paper ready for use. The **examples** show a suitable format for writing solutions, although in many cases the details of a calculation are omitted and should be worked through by the reader.

RED MEANS STOP! The in-chapter **exercises** are of a routine nature, and are designed to be done quickly, when first learning the material. You should always do these exercises and check your answers before proceeding. They are designed to make sure you have some understanding of the material before you proceed. The end of chapter problems are there for additional practice.

It is essential to know the definitions before you try to solve problems. Memorize the statements, but at the same time have a geometrical picture in mind, and then with time you will develop a clear understanding of the concepts.

Some parts of the notes (for example, Chapters 2, 3 and 5) are more theoretical, and hence more difficult, than other parts. It takes time and mathematical maturity to fully understand the fundamental concepts of limit, continuity and differentiability. However, there is no need to feel discouraged if you find these topics difficult, because you can still press on and obtain a working knowledge of multivariable calculus, as required for applications.

Understanding and writing proofs is usually the most difficult aspect of a course in mathematics. Of course it is possible to apply a theorem without knowing the proof. You just have to *believe* that it is true ("trust me ..."). However, there are long term benefits in studying the proofs of theorems even if you do not plan to become a mathematician. Firstly, you will repeatedly apply the definitions, and this will reinforce your understanding of the basic concepts, making it possible for you to apply these concepts elsewhere. Secondly, studying proofs is excellent training for the mind in logical thinking. This experience will benefit you in later years, most immediately in taking courses of a more theoretical nature (e.g. a theoretical course in Computer Science).

Students' grades in assignments, tests and the final examination will be influenced by how clearly the ideas are expressed, and by how well the solutions are organized. It is therefore important that students ensure that they understand not simply how to obtain an answer which is technically correct, but also how to present a cogent mathematical argument.

Acknowledgements

Thanks are expressed to:

David Siegel: for many discussions and suggestions which have substantially affected the structure and detailed contents of this text.

Mike La Croix: for the amazing figures which he created for the text, and for his assistance on editing, formatting and LaTex'ing.

Mark Soric: for LaTeX'ing.

Kathleen Wilkie: for proof-reading.

Joe Rocca: for proof-reading and checking solutions.

Stephen New, Nico Spronk, Lilia Krivodonova, Ronny Wan, and the many others who have contributed suggestions for the material and formatting.

Faisal Al Faisal, for making updates to the course notes for editions 6.2 and 7.0, which included reordering the latter part of the course so that the difficult topic of double integrals can be seen earlier.

Chapter 1

Graphs of Scalar Functions

1.1 Scalar Functions

One of the most important concepts in mathematics is that of a function. Recall that a function $f : A \to B$ associates with each element $a \in A$ a unique element $f(a) \in B$ called the **image** of a under f. The set A is called the **domain** of f and is denoted by $D(f)$. The set B is called the **codomain** of f. The subset of B consisting of all $f(a)$ is called the **range** of f and is denoted $R(f)$.

We will first extend what we did in single variable calculus to functions of several variables. We will usually look at real functions of two variables whose domain is a subset of $\mathbb{R}^2$ and whose codomain is $\mathbb{R}$. That is, we consider functions f which map points $(x, y) \in \mathbb{R}^2$ to a real scalar $f(x, y) \in \mathbb{R}$. We write $z = f(x, y)$. However, we will also consider more general functions $f(x_1, \dots, x_n)$ which map subsets of $\mathbb{R}^n$ to $\mathbb{R}$.

REMARK

Although strictly speaking, $f(x, y)$ denotes the value of the function f at the point (x, y), it is common practice to use the phrase "the function $f(x, y)$" to stress which independent variables the function is dependent on.

DEFINITION

Scalar Function

A **scalar function** $f(x_1, \dots, x_n)$ of n-variables is a function whose domain is a subset of $\mathbb{R}^n$ and whose range is a subset of $\mathbb{R}$.

REMARK

We will sometimes use $\mathbf{x}$ to represent a point in $\mathbb{R}^n$. Note that there will be several times in the course where it is convenient to view points in $\mathbb{R}^n$ as a vectors in $\mathbb{R}^n$ to make use of results from linear algebra.

EXAMPLE 1

Let $f: \mathbb{R}^2 \to \mathbb{R}$ be defined by $f(x, y) = 2x + 3y + 1$. Find $f(1, -4)$ and $f(1, 1)$.

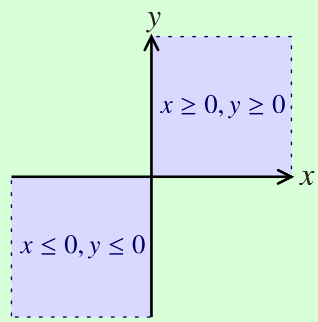
Solution: We have

$$f(1, -4) = 2(1) + 3(-4) + 1 = -9$$
$$f(1, 1) = 2(1) + 3(1) + 1 = 6$$

EXAMPLE 2

Find the largest subset of $\mathbb{R}^2$ that can serve as a domain for $f(x, y) = \sqrt{xy}$. Find the range of the resulting function f defined on this domain.

Solution: We cannot take the square root of a negative number, so we require $xy \geq 0$. Thus, the domain is the set $x \geq 0$, $y \geq 0$ and $x \leq 0$, $y \leq 0$. Since this is a subset of $\mathbb{R}^2$, it is easy to represent with a picture.



For the range, we notice that $f(x, y) = \sqrt{xy} \geq 0$. To see that the range of f contains all non-negative real numbers, observe that for any non-negative real number c we have that $f(c^2, 1) = \sqrt{c^2} = |c| = c$.

REMARK

Technically speaking, a function comes with a domain by definition. Thus in the preceding example we should have asked for the largest subset of $\mathbb{R}^2$ that can serve as a domain for the expression $f(x, y) = \sqrt{xy}$.

Our convention going forwards is that, given such an expression $f(x, y)$, we will tacitly assume (unless explicitly stated otherwise) that its domain is the largest possible subset in $\mathbb{R}^2$ where the expression is well-defined. Thus, when we ask for the *domain* of such an $f(x, y)$, we are referring to this aforementioned subset.

EXAMPLE 3

Find the domain and range of $g(x, y) = \dfrac{x^2 - y^2}{|x| + |y|}$.

Solution: Observe g is defined whenever $(x, y) \neq (0, 0)$. So, the domain is $\mathbb{R}^2 - \{(0, 0)\}$.

The range is a little more difficult to see. We need to determine all values we can get from g by taking points in our domain. We first consider points $(c, 0)$, $c \neq 0$. We get

$$g(c, 0) = \frac{c^2 - 0^2}{|c| + |0|} = |c|$$

Hence, g can take any positive value. Similarly, points $(0, d)$, $d \neq 0$ give

$$g(0, d) = \frac{0^2 - d^2}{|0| + |d|} = -|d|$$

Thus, g can also take any negative value. Finally, observe that $g(1, 1) = 0$. Therefore, the range of g is $\mathbb{R}$.

EXERCISE 1

Sketch the domain and find the range of the following functions:

(a) $f(x, y) = \ln(1 - x^2 - y^2)$.

(b) $g(x, y) = \sqrt{16 - x^2 + y^2}$.

For more complicated functions, it could be extremely difficult to determine their range. When we had such situations with single variable functions we often found it helpful to sketch the graph of the function. So, we now determine how to sketch a graph of a function $f(x, y)$.

1.2 Geometric Interpretation of $z = f(x, y)$

When we graph a function $y = f(x)$ we plot points $(a, f(a))$ in the xy-plane. Observe that we can think of $f(a)$ as representing the height of the graph $y = f(x)$ above (or below if negative) the x-axis at $x = a$.

We define the **graph** of a function $f(x, y)$ as the set of all points $(a, b, f(a, b))$ in $\mathbb{R}^3$ such that $(a, b) \in D(f)$. We think of $f(a, b)$ as representing the height of the graph $z = f(x, y)$ above the xy-plane at the point $(x, y) = (a, b)$.

EXAMPLE 1

Let f be defined by $f(x, y) = c_1x + c_2y + c_3$, where c_1, c_2, c_3 are real constants. We recognize this as the equation of a plane in $\mathbb{R}^3$. That is, the graph of $z = f(x, y)$ is a plane.

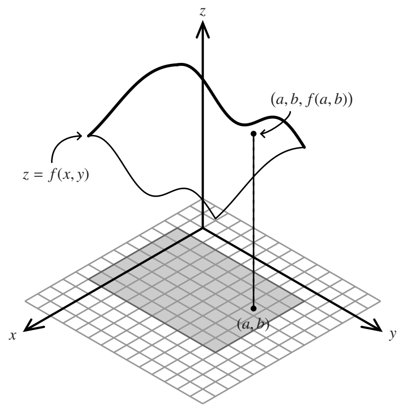
In general, surfaces $z = f(x, y)$ can be quite complicated. To help us visualize and/or sketch these surfaces, we look at 2-dimensional slices of the surface.

DEFINITION
Level Curves

The **level curves** of a function $f(x, y)$ are the curves

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

where k is a constant in the range of f.

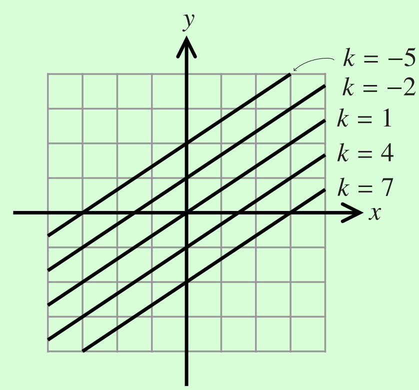


EXAMPLE 2 What are the level curves of the function defined by $f(x, y) = 2x - 3y + 1$?

Solution: We observe that $R(f) = \mathbb{R}$. So, the level curves of f are

$$2x - 3y + 1 = f(x, y) = k, \quad k \in \mathbb{R}$$

Sketching gives a family of parallel lines:



Observe that the level curve $f(x, y) = k$ is the intersection of $z = f(x, y)$ and the horizontal plane $z = k$. Thus, in our family of level curves, each value of k represents the height of that level curve above the xy-plane. For this reason, the family of level curves is often called a **contour map** or a **topographic** map.

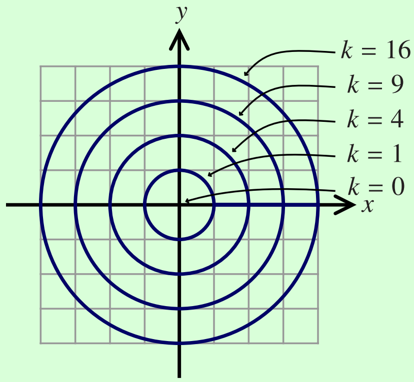
EXAMPLE 3 Consider the functions defined by

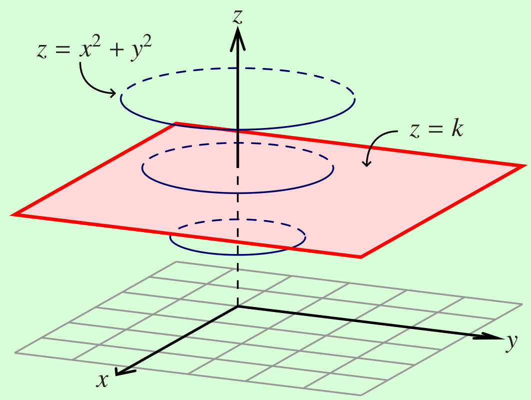
$$f(x, y) = x^2 + y^2, \qquad g(x, y) = x^2 - y^2, \qquad h(x, y) = x^2$$

(a) Sketch the level curves of f and use them to sketch the surface $z = f(x, y)$.

Solution: We first observe that $D(f) = \mathbb{R}^2$ and $R(f) = \{z \in \mathbb{R} \mid z \geq 0\}$ since $x^2 + y^2 \geq 0$. Hence, k can take on values $k \geq 0$. For $k > 0$, the level curves of f are the circles $x^2 + y^2 = k$ with center (0,0). However, notice that for $k = 0$, the level curve is $x^2 + y^2 = 0$ which is just the single point $(0, 0)$. This is called an **exceptional level curve**.

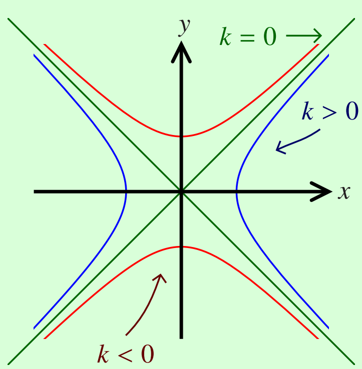
Remembering that k represents the height of the level curve $f(x, y) = k$ above the xy-plane, we sketch the surface by drawing the circles in the appropriate planes $z = k$ in $\mathbb{R}^3$. We get the surface below which is called a **paraboloid**.

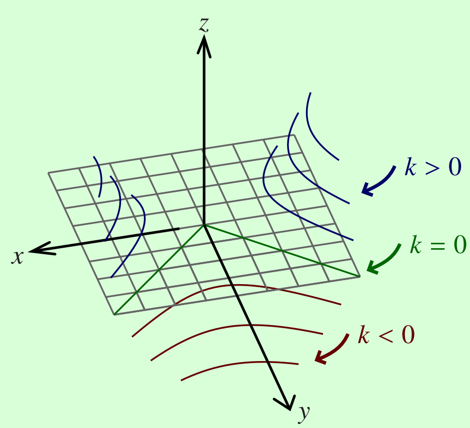




(b) Sketch the level curves of g and use them to sketch the surface $z = g(x, y)$.

Solution: We first observe that $D(g) = \mathbb{R}^2$ and $R(g) = \mathbb{R}$. For any $k \in \mathbb{R}$ we sketch the level curves $x^2 - y^2 = k$ which we recognize as a family of hyperbola with asymptotes $y = \pm x$ corresponding to $x^2 - y^2 = 0$. Using these to sketch the surface, we get a **saddle surface**.



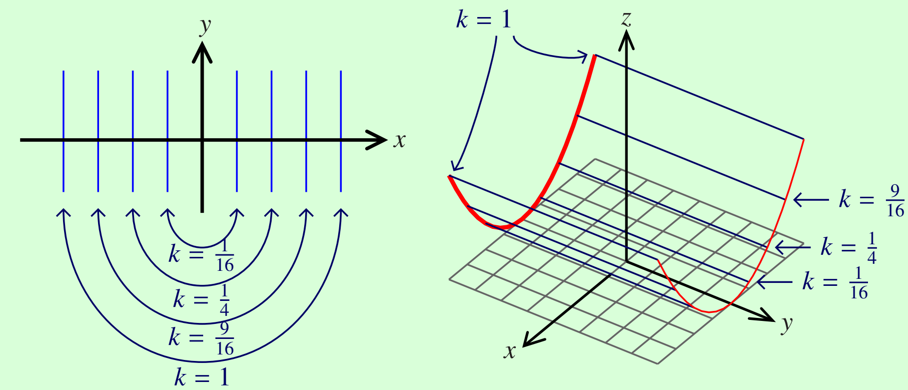


(c) Sketch the level curves of h and use them to sketch the surface $z = h(x, y)$.

Solution: We have that $D(h) = \mathbb{R}^2$ and $R(h) = \{z \in \mathbb{R} \mid z \geq 0\}$. Thus, for $k \geq 0$ we have level curves

$$x^2 = k \Rightarrow x = \pm\sqrt{k}$$

Hence, the level curves are pairs of vertical straight lines. Using these to sketch the surface, we get a **parabolic cylinder**.



REMARK

Level curves occur in everyday life. For example, the elevation of the earth's surface above sea level is described by an equation $z = h(x, y)$ where x is the latitude and y is the longitude of the position. A contour map shows the curves of constant elevation, $h(x, y) = k$, which are precisely the level curves of h.

Some other examples include use in weather maps to show curves of constant temperature called **isotherms**, in marine charts to indicate water depths, and in barometric pressure charts to show curves of constant pressure called **isobars**.

In general, it is not always possible to sketch the level curves of a given function $f(x, y)$ by inspection. Later in the course, we will develop some results which can be used to obtain information about the level curves of a function.

One can also obtain insight into the shape of a surface $z = f(x, y)$ by sketching the curves of intersection of the surface with other planes.

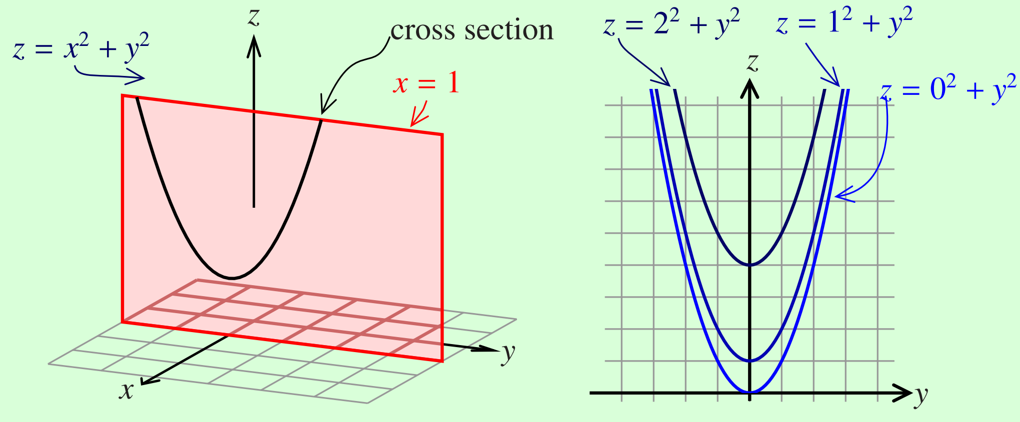
DEFINITION
Cross Sections

A **cross section** of a surface $z = f(x, y)$ is the intersection of $z = f(x, y)$ with a plane.

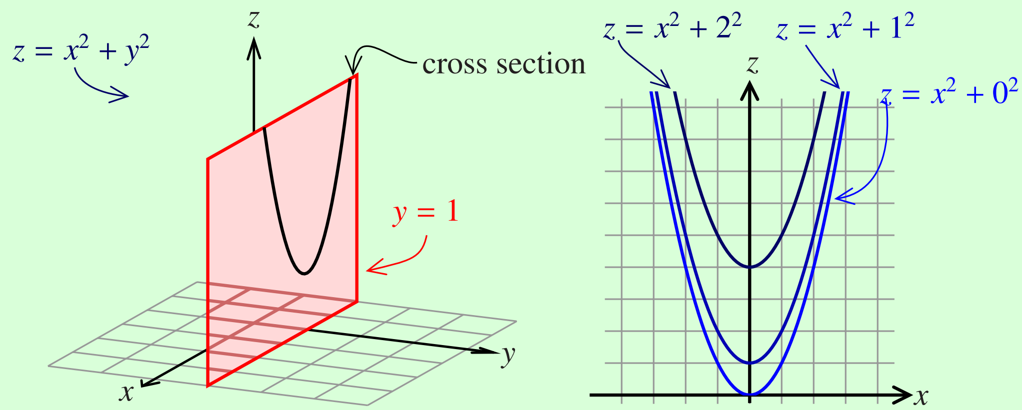
For the purpose of sketching the graph of a surface $z = f(x, y)$, it is useful to consider the cross sections formed by intersecting $z = f(x, y)$ with the vertical planes $x = c$ and $y = d$.

EXAMPLE 4

Let $f(x, y) = x^2 + y^2$. The cross sections formed by intersection $z = f(x, y)$ with $x = c$ for $c = 0, 1, 2$, are:



The cross sections formed by intersecting $z = f(x, y)$ with $y = d$ for $d = 0, 1, 2$ are:



REMARK

When sketching graphs, for simplicity, when we say to sketch the cross sections of a surface, we mean to sketch the family of cross sections $z = f(c, y)$ and $z = f(x, d)$ formed by intersection the surface with the vertical planes $x = c$ and $y = d$.

EXERCISE 1 Sketch the cross sections of $g(x, y) = x^2 - y^2$ and $h(x, y) = x^2$.

EXERCISE 2 Sketch the level curves and cross sections of $f(x, y) = \sqrt{x^2 + y^2}$ and use them to sketch the surface $z = f(x, y)$.

Generalization:

DEFINITION
Level Surfaces

A **level surface** of a scalar function $f(x, y, z)$ is defined by

$$f(x, y, z) = k, \quad k \in R(f)$$

EXAMPLE 5

The level surfaces of $f(x, y, z) = x^2 + y^2 + z^2$ are the family of spheres $x^2 + y^2 + z^2 = k$, for $k > 0$. In the exceptional case $k = 0$, the level surface is the single point $(0, 0, 0)$.

DEFINITION
Level Sets

A **level set** a scalar function $f(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^n$ is defined by

$$f(\mathbf{x}) = k, \quad k \in R(f)$$

EXAMPLE 6

Let f be defined by:

$$f(x_1, \dots, x_n) = x_1^2 + x_2^2 + \cdots + x_n^2$$

The level sets of $f(\mathbf{x}) = k$, $k > 0$ in $\mathbb{R}^n$ are called $(n-1)$**-spheres**, denoted by S^{n-1}. If $n = 3$, we obtain 2-spheres, S^2, as in Example 5.

Chapter 1 Problem Set

1. For the following functions $f(x, y)$, sketch typical level curves and any exceptional level curves. Level curves are defined by $f(x, y) = k$, where k is a constant.
   (a) $f(x, y) = 4x^2 - y^2$
   (b) $f(x, y) = x^2 + 4y^2 - 9$
   (c) $f(x, y) = x^2 + y^2 - 4(x + y)$
   (d) $f(x, y) = e^{4-x^2-y^2} - 1$
   (e) $f(x, y) = 2xy - y^2$

   Discuss the following:
   - What values can k assume? i.e. determine the range of f.
   - How do the level curves change as k increases?
   - Shade in any region of the xy-plane for which $k > 0$.
   - Sketch some typical cross-sections $x = c$, and some typical cross-sections $y = d$.
   - Describe/draw/visualize the surface $z = f(x, y)$ in 3-space.
2. Repeat question 1 for the following functions.
   (a) $f(x, y) = 1 - x^4 - y^4$
   (b) $f(x, y) = 1 - (x^2 + y^2 - 4)^2$
3. Let $f(x, y) = \sqrt{4 + x^2 - y^2}$.
   (a) Sketch the domain of f and state the range of f.
   (b) Sketch some level curves and cross sections $x = c$ and $y = d$ of f.
4. Let $f(x, y) = |1 - x^2 - y^2|$.
   (a) Sketch the domain of f and state the range of f.
   (b) Sketch some level curves and cross sections $x = c$ and $y = d$ of f.
5. Let $f(x, y) = \sqrt{x^2 - y^2}$.
   (a) Sketch the domain and state the range of f.
   (b) Sketch some level curves and cross sections $x = c$ and $y = d$ of f.
   (c) Print a Maple plot of the surface $z = \sqrt{x^2 - y^2}$. Draw the level curve $1 = \sqrt{x^2 - y^2}$ and the cross-section $z = \sqrt{x^2 - 0^2}$ on your print out.
6. For each of the given functions $f : \mathbb{R}^2 \to \mathbb{R}$,
   (i) sketch the domain of f and state the range of f.
   (ii) Sketch some level curves and cross sections $x = c$ and $y = d$ of f.

(iii) sketch the surface $z = f(x, y)$. Verify your answer using Maple. (If you plot the surface in Maple before attempting to sketch the surface, you are missing the point of the exercise).

(a) $f(x, y) = \sqrt{4 + x^2 - y^2}$

(b) $f(x, y) = x^2$

(c) $f(x, y) = \sqrt{|1 - x^2 - y^2|}$

7. * The temperature of a metal rod at position x, $0 \le x \le 1$, and at time t, $t \ge 0$ is given by $u(t, x) = 100e^{-t} \sin \pi x$. Sketch the level curves $u = 0, 25, 75, 100$. Shade the region of the tx-plane for which $u > 75$.

8. Suppose that the temperature $T(x, y)$ of an annular metal plate ($1 \le x^2 + y^2 \le e$) is given by

$$T(x, y) = 2\ln(x^2 + y^2) + 3.$$

Draw the annulus in the xy-plane, and sketch several isotherms (curves of constant temperature). Is the spacing of the curves uniform, or do they 'bunch up' near the inside or outside of the annulus?

9. Imagine a hill whose elevation z above sea-level (in meters) at position (x, y) is given by $z = f(x, y)$, where $f(x, y) = 1000 - 9x^2 - 4y^2$. A hiker, starting at position (10,5,0), walks up the hill in a south-westerly direction (the positive y-axis points northwards). Find the maximum elevation reached by the hiker.

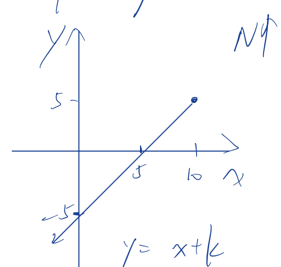
10. * A function g is defined by

$$g(x, y) = \int_x^y e^{-t^2}\, dt$$

Sketch the level curves of g.

Note: An asterisk(*) denotes a challenging problem.

9. $z = 1000 - 9x^2 - 4y^2$.



$y = x + k$

$5 = 10 + k$

$k = -5$.

$y = x - 5$.

$z = 1000 - 9x^2 - 4(x-5)^2$

$= 1000 - 9x^2 - 4(x^2 + 25 - 10x)$

$= 1000 - 13x^2 - 100 + 40x$

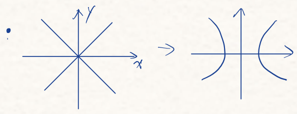
$= -13x^2 + 40x + 900$.

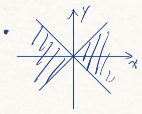
max $z = 900.31$.

1. For the following functions $f(x, y)$, sketch typical level curves and any exceptional level curves. Level curves are defined by $f(x, y) = k$, where k is a constant.
   (a) $f(x, y) = 4x^2 - y^2$
   (b) $f(x, y) = x^2 + 4y^2 - 9$
   (c) $f(x, y) = x^2 + y^2 - 4(x + y)$
   (d) $f(x, y) = e^{4-x^2-y^2} - 1$
   (e) $f(x, y) = 2xy - y^2$

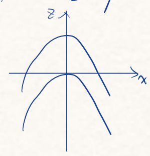
   Discuss the following:
   - What values can k assume? i.e. determine the range of f.
   - How do the level curves change as k increases?
   - Shade in any region of the xy-plane for which $k > 0$.
   - Sketch some typical cross-sections $x = c$, and some typical cross-sections $y = d$.
   - Describe/draw/visualize the surface $z = f(x, y)$ in 3-space.

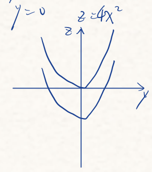
1. (a) • $k = 4x^2 - y^2$.
$k \in \mathbb{R}$.

• 

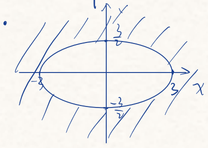
• 

• $x = \pm 1$ $z = 4 - y^2$ $y = \pm 1$ $z = 4x^2 - 1$
$x = 0$ $z = -y^2$ $y = 0$ $z = 4x^2$



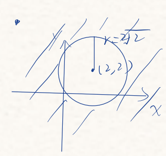


(b) • $k = x^2 + 4y^2 - 9$
$[-9, \infty)$
• $x^2 + 4y^2 = 9 + k$.
radius of oval increase

• 

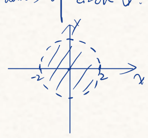
• 

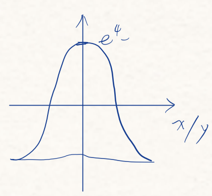
(c) • $k \in \mathbb{R}$
• $k = x^2 + y^2 - 4x - 4y$
$= (x^2 - 4x + 4) + (y^2 - 4y + 4) - 8$
$k + 8 = (x-2)^2 + (y-2)^2$
radius increase

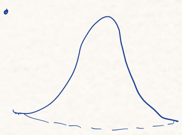
• 

• 

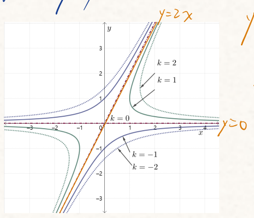
(d) • $(-1, e^4 - 1]$
• $k + 1 = e^{4-x^2-y^2}$
$4 - x^2 - y^2 = \ln(k+1)$
$x^2 + y^2 = 4 - \ln(k+1)$
$k\uparrow$. $\ln(k+1)\uparrow$. $4 - \ln(k+1)\downarrow$
radius of circle $\downarrow$.

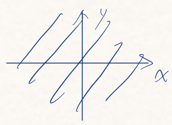
• 

• 

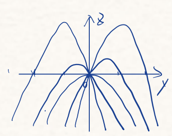
• 

(e) • $k = 2xy - y^2$ $k \in \mathbb{R}$.
当 $k=0$
$y(2x - y) = 0$
$y = 0$ $y = 2x$

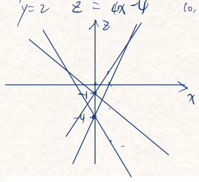
• 

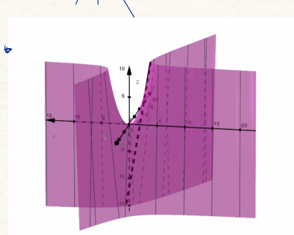
• 

• $x = -2$ $z = -4y - y^2 = -y(y+4)$
$x = -1$ $z = -2y - y^2 = -y(y+2)$
$x = 0$ $z = -y^2$
$x = 1$ $z = 2y - y^2 = -y(y-2)$
$x = 2$ $z = 4y - y^2 = -y(y-4)$



$y = -2$ $z = -4x - 4$ $(0,-4)$ $(1,-8)$
$y = -1$ $z = -2x - 1$ $(0,-1)$ $(1,-3)$
$y = 0$ $z = 0$
$y = 1$ $z = 2x - 1$ $(0,-1)$ $(1,1)$
$y = 2$ $z = 4x - 4$ $(0,-4$



• 

2. Repeat question 1 for the following functions.

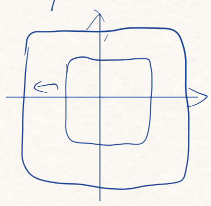
(a) $f(x, y) = 1 - x^4 - y^4$

(b) $f(x, y) = 1 - ((x^2 + y^2) - 4)^2$

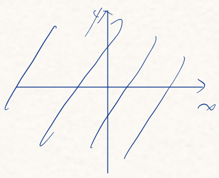
(a) • $(-\infty, 1]$

• $k = 1 - x^4 - y^4$.

$x^4 - y^4 = k - 1$



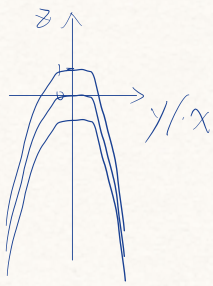
•



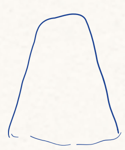
• $x = \pm 2 \quad z = 1 - 16 - y^4$

$x = \pm 1 \quad z = 1 - 1 - y^4$

$x = 0 \quad z = 1 - y^4$



•



(b) $z = 1 - [(x^2+y^2)^2 + 16 - 8(x^2+y^2)]$

$= 1 - (x^2+y^2)^2 - 16 + 8(x^2+y^2)$

$= -(x^2+y^2)^2 + 8(x^2+y^2) - 15$

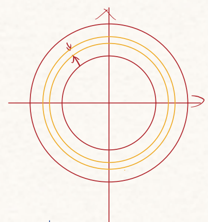
• $(-\infty, 1]$

$-x^2 + 8x - 15$

$x^2+y^2 = 4. \quad z = -16 + 32 - 15$

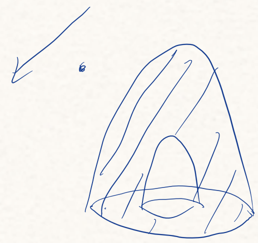
$= 1$

•



• 类似于 (a)

•



3. Let $f(x, y) = \sqrt{4 + x^2 - y^2}$.

(a) Sketch the domain of f and state the range of f.

(b) Sketch some level curves and cross sections $x = c$ and $y = d$ of f.

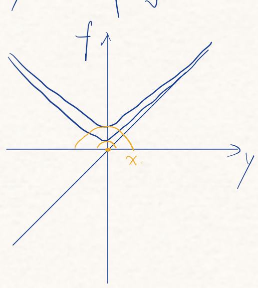
(a) $x \in R$. $y \in R$.

$f(x,y) \in [0, +\infty)$

(b) $y = 0 \quad f = \sqrt{4 + x^2}$

$y = \pm 1 \quad f = \sqrt{4 + x^2 - 1} = \sqrt{3 + x^2}$

$y = \pm 2 \quad f = \sqrt{4 + x^2 - 4} = \sqrt{x^2} = x$



$0 \sim \infty$

4. Let $f(x, y) = |1 - x^2 - y^2|$.

(a) Sketch the domain of f and state the range of f.

(b) Sketch some level curves and cross sections $x = c$ and $y = d$ of f.

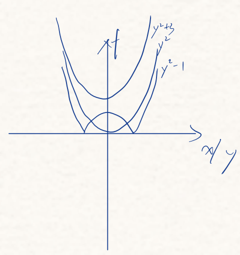
(a) $x \in R$. $y \in R$.

$f(x,y) \in [1, \infty)$

(b) $x = 0 \quad f = |1 - y^2| = |y^2 - 1|$

$x = \pm 1 \quad f = |-y^2| = y^2$

$x = \pm 2 \quad f = |-3 - y^2| = y^2 + 3$.



Chapter 2

Limits

2.1 Definition of a Limit

Recall for a function $f(x)$ we defined $\lim_{x \to a} f(x) = L$ to mean that the values of $f(x)$ can be made arbitrarily close to L by taking x sufficiently close to a. More precisely, for every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$|f(x) - L| < \epsilon \quad \text{whenever} \quad 0 < |x - a| < \delta$$

and $\lim_{x \to a} f(x) = L$ if and only if $\lim_{x \to a^-} f(x) = L = \lim_{x \to a^+} f(x)$.

Similarly, for scalar functions $f(x, y)$, we want $\lim_{(x,y) \to (a,b)} f(x, y) = L$ to mean the values of $f(x, y)$ can be made arbitrarily close to L by taking (x, y) sufficiently close to (a, b). For the one variable case we could only approach the limit from two directions, left and right. For multivariable scalar functions our domain is now multidimensional, so we can approach the limits from infinitely many directions. Moreover, we are not restricted to straight lines either; we can approach (a, b) along any smooth curve as well! Hence, to generalize the precise definition of a limit we need to generalize the concepts of an interval.

DEFINITION
Neighborhood

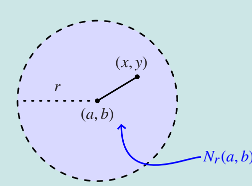
An ***r*-neighborhood** of a point $(a, b) \in \mathbb{R}^2$ is a set

$$N_r(a, b) = \{(x, y) \in \mathbb{R}^2 \mid \|(x, y) - (a, b)\| < r\}$$

REMARK

Recall that $\|(x, y) - (a, b)\|$ is the Euclidean distance in $\mathbb{R}^2$. That is,

$$\|(x, y) - (a, b)\| = \sqrt{(x - a)^2 + (y - b)^2}$$



Thus, we get:

DEFINITION

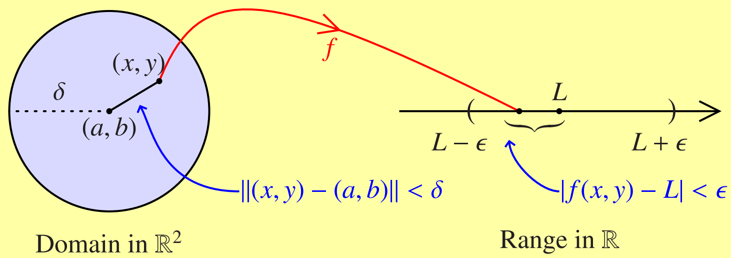
Limit

Assume $f(x,y)$ is defined in a neighborhood of (a,b), except possibly at (a,b). If for every $\epsilon > 0$ there exists a $\delta > 0$ such that

$$0 < \|(x,y)-(a,b)\| < \delta \quad \text{implies} \quad |f(x,y)-L| < \epsilon$$

then

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



Using the precise definition can be quite complicated even for relatively simple limits. Thus, we will instead use the definition to prove theorems to make finding limits easier.

2.2 Limit Theorems

In extending our definition of a limit to functions $f(x,y)$ we would hope that we have preserved all of our properties of limits we had for single variable functions (otherwise it would not be a very good generalization!).

THEOREM 1

If $\lim\limits_{(x,y)\to(a,b)} f(x,y)$ and $\lim\limits_{(x,y)\to(a,b)} g(x,y)$ both exist, then

(a) $\lim\limits_{(x,y)\to(a,b)} [f(x,y)+g(x,y)] = \lim\limits_{(x,y)\to(a,b)} f(x,y) + \lim\limits_{(x,y)\to(a,b)} g(x,y)$.

(b) $\lim\limits_{(x,y)\to(a,b)} [f(x,y)g(x,y)] = \left[\lim\limits_{(x,y)\to(a,b)} f(x,y)\right]\left[\lim\limits_{(x,y)\to(a,b)} g(x,y)\right]$.

(c) $\lim\limits_{(x,y)\to(a,b)} \dfrac{f(x,y)}{g(x,y)} = \dfrac{\lim\limits_{(x,y)\to(a,b)} f(x,y)}{\lim\limits_{(x,y)\to(a,b)} g(x,y)}$, provided $\lim\limits_{(x,y)\to(a,b)} g(x,y) \neq 0$.

Proof: We will prove (a) and leave (b) and (c) as exercises. Let $\epsilon > 0$. Since $\lim\limits_{(x,y)\to(a,b)} f(x,y) = L_1$ and $\lim\limits_{(x,y)\to(a,b)} g(x,y) = L_2$ both exist, by definition of a limit, there exists a $\delta > 0$ such that

$$0 < \|(x,y)-(a,b)\| < \delta \text{ implies } |f(x,y)-L_1| < \frac{1}{2}\epsilon \text{ and } |g(x,y)-L_2| < \frac{1}{2}\epsilon$$

Thus, if $0 < \|(x,y)-(a,b)\| < \delta$, then

$$\begin{aligned}\left|f(x,y)+g(x,y)-(L_1+L_2)\right| &= \left|[f(x,y)-L_1]+[g(x,y)-L_2]\right|\\ &\leq |f(x,y)-L_1| + |g(x,y)-L_2|\\ &< \frac{\epsilon}{2}+\frac{\epsilon}{2} = \epsilon\end{aligned}$$

as required. $\square$

THEOREM 2 If $\lim\limits_{(x,y)\to(a,b)} f(x,y)$ exists, then the limit is unique.

Proof: Assume that $\lim\limits_{(x,y)\to(a,b)} f(x,y) = L_1$ and $\lim\limits_{(x,y)\to(a,b)} f(x,y) = L_2$. Then,

$$L_1 - L_2 = \lim_{(x,y)\to(a,b)} f(x,y) - \lim_{(x,y)\to(a,b)} f(x,y) = \lim_{(x,y)\to(a,b)} [f(x,y)-f(x,y)] = 0$$

Hence, $L_1 = L_2$ and so the limit is unique. $\square$

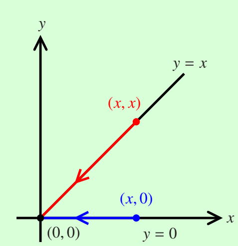
2.3 Proving a Limit Does Not Exist

Recall for a function of one variable, we often showed a limit did not exist by showing the left-hand limit did not equal the right-hand limit and using the fact that the limit is unique. For multivariable functions, we will essentially do the same thing, only now we have to remember that we are able to approach (a,b) along any smooth curve.

EXAMPLE 1 Let f be defined by $f(x,y) = \dfrac{xy}{x^2+y^2}$, for $(x,y) \neq (0,0)$. Prove that $\lim\limits_{(x,y)\to(0,0)} f(x,y)$ does not exist.

Solution: To prove this does not exist we just need to approach the limit along two paths that give different values.

We first approach the limit along the line $y = 0$. Notice that by holding y constant, we are turning this limit of a function of two variables into a limit of a function of a single variable x. We get



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

Now, approach the limit along the line $y = x$. This again changes the limit of a function of two variables into the limit of a function of one variable. We get

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

Since $f(x,y)$ approaches different values as (x,y) tends to $(0,0)$ along different paths, the limit does not exist.

We often can approach the limit along infinitely many lines or smooth curves at the same time by introducing an arbitrary coefficient m. If the value of the limit depends on the value of m, then it is not unique and hence, the limit does not exist.

EXAMPLE 2 Prove that $\displaystyle\lim_{(x,y)\to(0,0)} \frac{\sin(xy)}{x^2+y^2}$ does not exist.

Solution: Approaching the limit along lines $y = mx$ we get

$$\begin{aligned}
\lim_{(x,y)\to(0,0)} \frac{\sin(x(mx))}{x^2+(mx)^2} &= \lim_{x\to 0} \frac{\sin(mx^2)}{x^2(1+m^2)} \\
&= \lim_{x\to 0} \frac{2mx\cos(mx^2)}{2x(1+m^2)} && \text{by L'Hôpital's rule} \\
&= \lim_{x\to 0} \frac{m\cos(mx^2)}{1+m^2} \\
&= \frac{m}{1+m^2}
\end{aligned}$$

Since the limit depends on m we can get different limits along different lines $y = mx$ and hence $\displaystyle\lim_{(x,y)\to(0,0)} \frac{\sin(xy)}{x^2+y^2}$ does not exist.

EXERCISE 1 Let $f(x,y) = \dfrac{|x|}{|x|+y^2}$, for $(x,y) \neq (0,0)$. Show that

$$\lim_{(x,y)\to(0,0)} f(x, mx) = 1$$

for all $m \in \mathbb{R}$, but $\displaystyle\lim_{(x,y)\to(0,0)} f(x,y)$ does not exist.

Hint: $y = mx$ *does not* describe all lines through the origin.

→ approach limit along $y=mx$.

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

$x>0$ $\displaystyle\lim_{x\to 0^+} \frac{|x|}{|x|+m^2x^2} = \lim_{x\to 0^+} \frac{x}{x+m^2x^2} = \lim_{x\to 0^+} x \frac{1}{1+m^2x} = 1$

$x<0$ $\displaystyle\lim_{x\to 0^-} \frac{|x|}{|x|+m^2x^2} = \lim_{x\to 0^-} \frac{-x}{-x+m^2x^2} = \lim_{x\to 0^-} (-x) \frac{1}{1-m^2x} = 1$

$$\lim_{(x,y)\to(0,0)} f(x, mx) = \lim_{(x,y)\to(0,0)} \frac{|x|}{|x|+y^2} = 1. \quad \text{QED}$$

→ approach limit along $x=0$.

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

$0 \neq 1$

$\therefore$ limit DNE.

EXAMPLE 3 Let $f(x,y) = \dfrac{x^2y}{x^4+y^2}$, for $(x,y) \neq (0,0)$. Show that $\lim\limits_{(x,y)\to(0,0)} f(x,y)$ does not exist.

Solution: As before we first test the limit along lines $y = mx$. We get

$$\lim_{(x,y)\to(0,0)} f(x,mx) = \lim_{x\to 0} \frac{x^2(mx)}{x^4+(mx)^2} = \lim_{x\to 0} \frac{mx}{x^2+m^2} = 0$$

and

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

These all give the same value, so we start testing curves. Of course, we don't want to start randomly guessing curves. To get a limit other than 0, we need the power of x everywhere in the denominator to match the power of x in the numerator (so that they cancel out). This prompts us to try the limit along $y = x^2$. We get

$$\lim_{(x,y)\to(0,0)} f(x,x^2) = \lim_{x\to 0} \frac{x^2(x^2)}{x^4+(x^2)^2} = \lim_{x\to 0} \frac{1}{2} = \frac{1}{2}$$

Since we have two different values along two different paths, the limit does not exist.

REMARK

1. We could have done the last example more efficiently by just testing $y = mx^2$ to begin with and showing the limit depends on m.

2. Make sure that all lines or curves you use actually approach the limit. A common error is to approach a limit like in Example 3 along a line $x = 1$... which of course is meaningless as it does not pass through $(0,0)$.

3. Example 3 shows that no matter how many lines and/or curves you test, you cannot use this method to prove a limit exists. Just because you haven't found two paths that give different values does not mean there isn't one!

EXERCISE 2 Prove that $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^3y}{x^6+y^2}$ does not exist.

EXERCISE 3 Prove that $\lim\limits_{(x,y)\to(1,0)} \dfrac{(x-1)(y+1)}{|x-1|+y}$ does not exist.

2.4 Proving a Limit Exists

Since we cannot use the previous methods to prove a limit exists, we prove another theorem to help us.

THEOREM 1 **(Squeeze Theorem)**

If there exists a function $B(x, y)$ such that

$$|f(x, y) - L| \leq B(x, y), \qquad \text{for all } (x, y) \neq (a, b)$$

in some neighborhood of (a, b) and $\lim\limits_{(x,y)\to(a,b)} B(x, y) = 0$, then

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

Proof: Let $\epsilon > 0$. Since $\lim\limits_{(x,y)\to(a,b)} B(x, y) = 0$ we have that there exists a $\delta > 0$ such that

$$0 < \|(x, y) - (a, b)\| < \delta \text{ implies } |B(x, y) - 0| < \epsilon$$

Hence, if $0 < \|(x, y) - (a, b)\| < \delta$, then we have

$$|f(x, y) - L| \leq B(x, y) = |B(x, y)| < \epsilon$$

as our hypothesis requires that $B(x, y) \geq 0$ for all $(x, y) \neq (a, b)$ in the neighborhood of (a, b). Therefore, by definition of a limit, we have

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

$\square$

EXERCISE 1 Our statement of the Squeeze Theorem above is not a direct generalization of the Squeeze Theorem we used in single variable calculus. What would the direct generalization of the Squeeze Theorem be? Show how your generalization and the theorem above are related.

EXAMPLE 1 Prove that $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^2 y}{x^2 + y^2} = 0$.

Solution: We have $f(x, y) = \dfrac{x^2 y}{x^2 + y^2}$ and $L = 0$. For $(x, y) \neq (0, 0)$ we obtain

$$|f(x, y) - L| = \left| \frac{x^2 y}{x^2 + y^2} - 0 \right| = \frac{x^2 |y|}{x^2 + y^2}$$

Since $y^2 \geq 0$, it follows that $x^2 \leq x^2 + y^2$, and hence

$$\frac{x^2|y|}{x^2+y^2} \leq (x^2+y^2)\frac{|y|}{x^2+y^2} = |y|$$

Thus,

$$0 \leq |f(x,y) - L| \leq |y|, \quad \text{for all } (x,y) \neq (0,0)$$

So, we are taking $B(x,y) = |y|$. By inspection,

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

Consequently, the Squeeze Theorem implies that

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

The next example illustrates some manipulations with inequalities.

EXAMPLE 2 Prove that

$$\frac{|2x^2-y^2|}{|x|+|y|} \leq 2|x| + |y|, \quad \text{for all } (x,y) \neq (0,0)$$

Solution: We can give a quick proof of this by realizing that the given inequality is equivalent to

$$|2x^2 - y^2| \leq (2|x| + |y|)(|x| + |y|) = 2|x|^2 + 3|x||y| + |y|^2, \quad \text{for all } (x,y) \neq (0,0)$$

The Triangle Inequality immediately gives us that $|2x^2 - y^2| \leq 2|x|^2 + |y|^2$ which is certainly $\leq 2|x|^2 + 3|x||y| + |y|^2$. This proves the inequality above, and hence the given inequality since the two are equivalent.

We will give a different proof that relies on manipulating the right-side of the original inequality. The reason being: in practice – like when using the Squeeze Theorem – we often have to work with functions where we are not *handed* an upper bound and instead must *discover* a suitable one by ourselves. The idea will be to manipulate the numerator so as to create a factor of $|x| + |y|$, which will cancel the denominator. For arbitrary (x,y), consider

$$\begin{aligned} |2x^2 - y^2| &= |2x^2 + (-y^2)| \\ &\leq |2x^2| + |-y^2|, \quad \text{by the Triangle Inequality} \\ &= 2|x|^2 + |y|^2 \end{aligned}$$

Since $|x| \leq |x| + |y|$, and $|y| \leq |x| + |y|$, we obtain

$$\begin{aligned} 2|x|^2 + |y|^2 &\leq 2|x|\big(|x| + |y|\big) + |y|\big(|x| + |y|\big) \\ &= \big(2|x| + |y|\big)\big(|x| + |y|\big) \end{aligned}$$

Hence,

$$\begin{aligned}\frac{|2x^2 - y^2|}{|x| + |y|} &\le \frac{(2|x| + |y|)(|x| + |y|)}{|x| + |y|} \\ &= 2|x| + |y|\end{aligned}$$

as required.

REMARK

Be careful when working with inequalities! For example, the statement

$$x < x^2$$

is **false** if $|x| < 1$. The appendix at the end of this chapter gives a brief review of inequalities.

EXERCISE 2 Prove that

$$\frac{|x^3 - y^3|}{x^2 + y^2} \le |x| + |y| \quad \text{for all} \quad (x, y) \neq (0, 0)$$

Does equality ever hold?

Before one can apply the Squeeze Theorem, one must have a possible limiting value L in mind. Of course, if you are asked to

$$\text{“Prove that } \lim_{(x,y)\to(a,b)} f(x, y) = L\text{”}$$

you are given the limiting value L, and can apply the Squeeze Theorem directly as in Example 1. On the other hand, if you are asked to

$$\text{“Determine whether } \lim_{(x,y)\to(a,b)} f(x, y) \text{ exists, and if so find its value”}$$

you should begin by letting (x, y) approach (a, b) along straight lines of different slope.

If the limiting value of $f(x, y)$ depends on the slope, then $\displaystyle\lim_{(x,y)\to(a,b)} f(x, y)$ does not exist.

If the limiting value of $f(x, y)$ does not depend on the slope and equals L, say, then $\displaystyle\lim_{(x,y)\to(a,b)} f(x, y)$ may exist and if it does exist, it equals L.

You should then try to apply the Squeeze Theorem to prove that the limit does exist and equals L.

If you fail to derive a suitable inequality, you cannot draw a conclusion, and you are faced with a dilemma . . .

perhaps a suitable inequality does exist, but you were not skillful enough to derive it,

OR

perhaps if you let (x, y) approach (a, b) along curves, then the you may get a limiting value other than L along one of those curves, in which case $\lim\limits_{(x,y)\to(a,b)} f(x, y)$ does not exist.

This can be a process of trial and error, but experience will help to shorten the process.

Here is an example:

EXAMPLE 3 Determine whether $\lim\limits_{(x,y)\to(0,0)} \dfrac{x^2 - |x| - |y|}{|x| + |y|}$ exists, and if so find its value.

Solution: Trying lines $y = mx$ we get

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

Since the value along each line is $L = -1$, we try to prove the limit is -1 with the Squeeze Theorem. Thus, we consider

$$\begin{aligned}\left|\frac{x^2 - |x| - |y|}{|x| + |y|} - (-1)\right| &= \left|\frac{x^2 - |x| - |y|}{|x| + |y|} + \frac{|x| + |y|}{|x| + |y|}\right| \\ &= \frac{x^2}{|x| + |y|} \\ &= \frac{|x| \cdot |x|}{|x| + |y|} \\ &\leq \frac{|x|(|x| + |y|)}{|x| + |y|} = |x|, \qquad \text{since } |x| \leq (|x| + |y|)\end{aligned}$$

Since $\lim\limits_{(x,y)\to(0,0)} |x| = 0$ we get $\lim\limits_{(x,y)\to(0,0)} \frac{x^2 - |x| - |y|}{|x| + |y|} = -1$ by the Squeeze Theorem.

EXERCISE 3 Consider f defined by

$$f(x, y) = \frac{x^2(x-1) - y^2}{x^2 + y^2}, \qquad \text{for } (x, y) \neq (0, 0)$$

Determine whether $\lim\limits_{(x,y)\to(0,0)} f(x, y)$ exists, and if so find its value.

REMARK

The concept of a neighbourhood, the definition of a limit, the Squeeze Theorem and the limit theorems are all valid for scalar functions $f(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^n$. In fact, to generalize these concepts, one only needs to recall that if $\mathbf{x} = (x_1, \dots, x_n)$ and $\mathbf{a} = (a_1, \dots, a_n)$ are in $\mathbb{R}^n$, then the Euclidean distance from $\mathbf{x}$ to $\mathbf{a}$ is

$$\|\mathbf{x} - \mathbf{a}\| = \sqrt{(x_1 - a_1)^2 + \cdots + (x_n - a_n)^2}$$

2.5 Appendix: Inequalities

The following statements can be taken as axioms (i.e. assumed properties) which define the notion of "less than" (denoted "<") for real numbers.[1]

Trichotomy Property: For any real numbers a and b, one and only one of the following holds:

$$a = b, \qquad a < b, \qquad b < a$$

Transitivity Property: If $a < b$ and $b < c$, then $a < c$.

Addition Property: If $a < b$, then for all c, $a + c < b + c$.

Multiplication Property: If $a < b$ and $c < 0$, then $bc < ac$.

Using these properties one can deduce other results.

The **absolute value** of a real number a is defined by

$$|a| = \begin{cases} a & \text{if } a \geq 0 \\ -a & \text{if } a < 0 \end{cases}$$

Three frequently used results, which follow from the axioms, are listed below.

1. $|a| = \sqrt{a^2}$.
2. $|a| < b$ if and only if $-b < a < b$.
3. the Triangle Inequality: $|a + b| \leq |a| + |b|$ for all $a, b \in \mathbb{R}$.

[1]One can equivalently use the notion of "greater than" (denoted ">"). The statement "$a > b$" means "$b < a$".

REMARK

When using the Squeeze Theorem, the most commonly used inequalities are:

1. the Triangle Inequality
2. if $c > 0$, then $a < a + c$
3. the cosine inequality $2|x||y| \leq x^2 + y^2$

One particularly common use of (2) is for things like

$$|x| = \sqrt{x^2} \leq \sqrt{x^2 + y^2}$$

We again stress that it is very important that one be careful when working with inequalities. Another common mistake is

$$\text{if } c > 0, \text{ then } |a + c| \leq a + c$$

Give an example to show where this statement is false.

Chapter 2 Problem Set

1. Prove that $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^4 - y^5}{x^4 + y^4}$ does not exist.

2. Prove that $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^2(y-3) - 6y^2}{x^2 + 2y^2} = -3.$

3. Find the limit, if it exists, or show that the limit does not exist.

   (a) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^4 + 2y^4}{x^4 + y^4}$

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

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

   (d) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^4 y^3}{x^8 + y^6}$

4. Find the limit, if it exists, or show that the limit does not exist.

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

   (b) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^3 y^2}{x^4 + y^6}$

5. Find the limit, if it exists, or show that the limit does not exist.

   (a) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{|x| - |y|}{|x| + |y|}$

   (b) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^2 - 3|x| - 3|y|}{|x| + |y|}$

   (c) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^5 y^2}{x^{10} + y^4}$

   (d) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^3 y^4}{x^6 + y^6}$

6. (a) Prove that $\sqrt{x^4 + y^4} \le x^2 + y^2$ for all $(x, y) \in \mathbb{R}^2$.

   (b) Determine whether $\displaystyle\lim_{(x,y)\to(0,0)} \frac{\sqrt{x^4 + y^4}}{\sqrt{x^2 + y^2}}$ exists.

7. Determine the values of p for which the following limit does or does not exist:

呃

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

8. * Let $f(x, y) = \dfrac{|x|^a |y|^b}{|x|^c + |y|^d}$ where a, b, c and d are positive numbers.

   (a) Prove that if $\dfrac{a}{c} + \dfrac{b}{d} > 1$ then $\displaystyle\lim_{(x,y)\to(0,0)} f(x, y)$ exists and equals zero.

   (b) Prove that if $\dfrac{a}{c} + \dfrac{b}{d} \le 1$, then $\displaystyle\lim_{(x,y)\to(0,0)} f(x, y)$ does not exist.

1. Prove that $\lim_{(x,y)\to(0,0)} \dfrac{x^4 - y^5}{x^4 + y^4}$ does not exist.

approach limit along $y = mx$

$$\lim_{(x,y)\to(0,0)} \frac{x^4 - m^5x^5}{x^4 + m^4x^4}$$

$$= \lim_{x\to 0} \frac{x^4(1-\cancel{m^5}\ mx)}{x^4(1+m^4)}$$

$$= \frac{1-\cancel{m^3}}{1+m^4} \qquad \frac{1}{1+m^4}$$

$\because$ limit depend on m

we get different limit along diff $y=mx$

$\therefore$ limit DNE

2. Prove that $\lim_{(x,y)\to(0,0)} \dfrac{x^2(y-3) - 6y^2}{x^2 + 2y^2} = -3.$

$f(x,y) = \dfrac{x^2(y-3)-6y^2}{x^2+2y^2} \qquad L=-3.$

$\to \because (x,y) \neq (0,0)$

$$\therefore |f(x,y) - L| = \left|\frac{x^2(y-3)-6y^2}{x^2+2y^2} + 3\right|$$

$$= \left|\frac{x^2(y-3)-6y^2+3(x^2+2y^2)}{x^2+2y^2}\right|$$

$$= \left|\frac{x^2(y-3+3)+\cancel{6y^2}-\cancel{6y^2}}{x^2+2y^2}\right|$$

$$= \left|\frac{x^2y}{x^2+2y^2}\right|$$

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

$\to \because 2y^2 \geq 0 \qquad \therefore x^2 \leq x^2+2y^2$

$$\therefore \frac{x^2|y|}{x^2+2y^2} \leq \frac{(x^2+2y^2)|y|}{x^2+2y^2} = |y|.$$

$\forall (x,y)\neq(0,0) \quad 0 \leq |f(x,y)-L| \leq |y|$

$\to$ take $B(x,y) = |y|$ $\quad \lim_{(x,y)\to(0,0)} |y| = 0$

$\therefore$ by squeeze theorem $\lim f = 0$

3. Find the limit, if it exists, or show that the limit does not exist.

(a) $\lim_{(x,y)\to(0,0)} \dfrac{x^4 + 2y^4}{x^4 + y^4}$ DNE

(b) $\lim_{(x,y)\to(0,0)} \dfrac{(x-y)^2}{|x| + |y|}$ $=0$

(c) $\lim_{(x,y)\to(0,0)} \dfrac{x^2 - 2|x| - |y|}{2|x| + |y|}$ ~~DNE~~ -1

(d) $\lim_{(x,y)\to(0,0)} \dfrac{x^4y^3}{x^8 + y^6}$ $\cancel{\tfrac{1}{2}}$ DNE

(a) $\to$ approach limit along $y=0$

$$\lim_{y\to 0} \frac{2y^4}{y^4} = 2$$

$\to$ approach limit along $x=0$

$$\lim_{x\to 0} \frac{x^4}{x^4} = 1$$

$2 \neq 1 \quad \therefore$ DNE

(b) $\to$ approach limit along $y=mx$

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

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

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

$= 0$

candidate limit $L=0$

$\to$ let $f(x) = \dfrac{(x-y)^2}{|x|+|y|}$

$$|f(x)-L| = \frac{(x-y)^2}{|x|+|y|} \leq \frac{(x-y)^2}{|x|}$$

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

$0 \leq |f(x)-L| \leq 0$

$\therefore$ by squeeze theorem. $\lim = 0$

(c) → approach limit along $y=mx$

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

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

$$= \lim_{x\to 0} \frac{|x|-2-|m|}{2+|m|}$$

$$= \frac{-2-|m|}{2+|m|} = -1$$

~~∴ limit depend on m.~~

~~∴ limit DNE~~

candidate limit : $L=-1$

→ let $f(x,y) = \frac{x^2-2|x|-|y|}{2|x|+|y|}$ $L=-1$

$$|f(x,y)-L| = \left|\frac{x^2-2|x|-|y|+2|x|+|y|}{2|x|+|y|}\right|$$

$$= \left|\frac{x^2}{2|x|+|y|}\right|$$

$$\lim_{(x,y)\to(0,0)} \left|\frac{x^2}{2|x|+|y|}\right| = \lim_{x\to 0} \frac{|x|}{2} = 0$$

∴ by squeeze theorem, lim = −1

(d) → approach limit along $y = x^{\frac{4}{3}}$

$$\lim_{(x,y)\to(0,0)} \frac{x^4\cdot(x^{\frac{4}{3}})^3}{x^8+(x^{\frac{4}{3}})^6}$$

$$= \lim_{x\to 0} \frac{x^8}{x^8+x^8}$$

$$= \frac{1}{2}$$

→ approach limit along $y=mx$

$$\lim_{x\to 0} \frac{x^4\cdot(m^3x^3)}{x^8+m^6x^6} = \lim_{x\to 0} \frac{x^6(m^3x)}{x^6(x^2+m^6)}$$

$$= \lim_{x\to 0} \frac{m^3x}{x^2+m^6} = 0.$$

→ ∴ $0 \neq \frac{1}{2}$

∴ DNE.

4. Find the limit, if it exists, or show that the limit does not exist.

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

(b) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^3y^2}{x^4+y^6}$

(a) → approach though $y=mx$.

$$\lim_{x\to 0} \frac{x^2-2|x|-2|mx|}{2|x|+|mx|}$$

$$= \lim_{x\to 0} \frac{|x|(|x|-2-2|m|)}{|x|(2+|m|)}$$

$$= \lim_{x\to 0} \frac{|x|-2-2|m|}{2+|m|}$$

$$= \frac{-2-2|m|}{2+|m|}$$

depend on $|m|$ → DNE

(b) → approach though $y = x^{\frac{2}{3}}$

$$\lim_{x\to 0} \frac{x^3\cdot(x^{\frac{2}{3}})^2}{x^4+(x^{\frac{2}{3}})^6}$$

$$= \lim_{x\to 0} \frac{x^3\cdot x^{\frac{4}{3}}}{x^4+x^4}$$

$$= \lim_{x\to 0} \frac{\cancel{x^4}\cdot x^{\frac{1}{3}}}{2\cancel{x^4}}$$

$$= 0$$

→ approach though $y=mx$.

$$\lim_{x\to 0} \frac{x^3\cdot m^2x^2}{x^4+m^6x^6}$$

$$= \lim_{x\to 0} \frac{x^4(m^2x)}{x^4(1+m^6x^2)}$$

$$= \lim_{x\to 0} \frac{m^2x}{1+m^6x^2}$$

$$= \lim_{x\to 0} \frac{2m}{2m^6x}.$$

$$= \infty$$

$0 \neq \infty$ DNE.

6. (a) Prove that $\sqrt{x^4 + y^4} \leq x^2 + y^2$ for all $(x, y) \in \mathbb{R}^2$.

(b) Determine whether $\displaystyle\lim_{(x,y)\to(0,0)} \frac{\sqrt{x^4 + y^4}}{\sqrt{x^2 + y^2}}$ exists.

(a) prove by contradiction $: \cdots > : \cdots$

$$\left(\sqrt{x^4+y^4}\right)^2 > (x^2+y^2)^2$$

$$x^4+y^4 > x^4+y^4+2x^2y^2.$$

$$2x^2y^2 < 0.$$

$\because 2x^2y^2 \geq 0$ contradicts to <0

$\therefore$ QED

(b) $\because \sqrt{x^4+y^4} \leq x^2+y^2.$

$$\therefore 0 \leq \frac{\sqrt{x^4+y^4}}{\sqrt{x^2+y^2}} \leq \frac{x^2+y^2}{\sqrt{x^2+y^2}}.$$

approach $\displaystyle\lim_{(0,0)} \frac{x^2+y^2}{\sqrt{x^2+y^2}}.$ along $y=mx.$

$$\lim_{(0,0)} \frac{x^2+m^2x^2}{\sqrt{x^2+m^2x^2}} = \lim_{(0,0)} \frac{x^2(1+m^2)}{x\sqrt{1+m^2}}$$

$$= \lim_{(0,0)} x \frac{(1+m^2)}{\sqrt{1+m^2}}$$

$$= 0$$

$\therefore$ by squeeze theorem.

$$\lim_{(0,0)} \frac{\sqrt{x^4+y^4}}{\sqrt{x^2+y^2}} = 0$$

7. Determine the values of p for which the following limit does or does not exist:

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

approach limit along $y=mx.$

$$\lim_{(x,y)\to(0,0)} \frac{|x||mx|^p}{x^2+m^2x^2}$$

$$\cancel{= \lim_{x\to 0} \frac{|x|(1+|m|^p)}{|x||x|(1+m^2)}} \times$$

$$\cancel{= \lim_{x\to 0} \frac{(1+|m|^p)}{|x|(1+m^2)}}$$

$$= \lim_{x\to 0} \frac{|x|^2|m|^p|x|^{p-1}}{|x|^2(1+m^2)}$$

$$= \lim_{x\to 0} \frac{|m|^p|x|^{p-1}}{1+m^2}$$

$$= \lim_{x\to 0} \frac{|m|^p}{1+m^2}|x|^{p-1}$$

limit only exist when $p-1>0.$

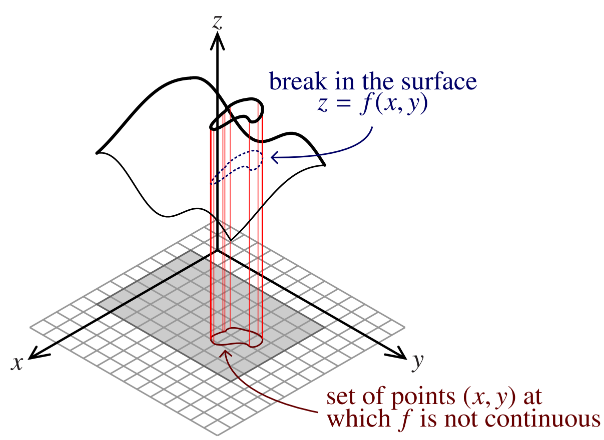
$p>1$

Chapter 3

Continuous Functions

3.1 Definition of a Continuous Function

In many situations, we shall require that a function $f(x, y)$ is **continuous**. Intuitively, this means that the graph of f (the surface $z = f(x, y)$) has no "breaks" or "holes" in it. As with functions of one variable, continuity is defined by using limits.



EXERCISE 1 Review the definition of a continuous function of one variable in your first year calculus text. Give an example (formula and graph) of a function $y = f(x)$ which is defined for all $x \in \mathbb{R}$, but is not continuous at $x = 1$.

Here is the formal definition:

DEFINITION
Continuous

A function $f(x, y)$ is **continuous** at (a, b) if and only if

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

Additionally, if f is continuous at every point in a set $D \subset \mathbb{R}^2$, then we say that f is continuous on D.

REMARK

There are really three requirements in this definition:

1. $\lim\limits_{(x,y)\to(a,b)} f(x, y)$ exists,
2. f is defined at (a, b),
3. the stated equality.

EXAMPLE 1 Let f be defined by

$$f(x, y) = \begin{cases} \dfrac{x^2 y}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{if } (x, y) = (0, 0) \end{cases}$$

Determine whether f is continuous at $(0, 0)$.

Solution: According to the definition we have to determine whether

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

This limit was established in Example 2.4.1. It follows that f is continuous at $(0, 0)$.

EXAMPLE 2 Prove that $f(x, y) = \begin{cases} \dfrac{xy}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{if } (x, y) = (0, 0) \end{cases}$ is not continuous at $(0, 0)$.

Solution: To prove that f is not continuous at $(0, 0)$, we just need to prove that the limit

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

does not equal 0. Therefore, if we can find one path such that the limit does not equal 0, then, since the value of a limit must be unique, this will prove that the limit cannot be equal to 0.

Approaching the limit along the line $y = x$ gives

$$\lim_{x\to 0} \frac{x^2}{x^2 + x^2} = \frac{1}{2}$$

Thus, the limit cannot equal 0, so the f is not continuous at $(0, 0)$.

EXAMPLE 3

Consider f defined by

$$f(x,y) = \frac{\sin(xy)}{x^2+y^2}, \quad \text{if } (x,y) \neq (0,0)$$

Can f be defined at $(0,0)$ so that the resulting function, whose domain is $\mathbb{R}^2$, is continuous at $(0,0)$?

Solution: By definition of continuity, we must determine whether

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

exists. It was shown in Example 2 in Section 2.3 that this limit does not exist. Thus, no matter what value we assign to $f(0,0)$ the resulting function will not be continuous at $(0,0)$.

EXERCISE 2

Let f be defined by $f(x,y) = \begin{cases} \dfrac{xy}{|x|+|y|} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$

Determine whether f is continuous at $(0,0)$.

3.2 The Continuity Theorems

One can often quickly prove that a function is continuous by applying certain theorems. The idea is to view a given function as being formed from simple functions by certain basic operations, which we now define.

DEFINITION

Operations on Functions

If $f(x,y)$ and $g(x,y)$ are scalar functions and $(x,y) \in D(f) \cap D(g)$, then:

1. the **sum** $f+g$ is defined by

$$(f+g)(x,y) = f(x,y) + g(x,y)$$

2. the **product** fg is defined by

$$(fg)(x,y) = f(x,y)g(x,y)$$

3. the **quotient** $\dfrac{f}{g}$ is defined by

$$\left(\frac{f}{g}\right)(x,y) = \frac{f(x,y)}{g(x,y)}, \quad \text{if } g(x,y) \neq 0$$

DEFINITION

Composite Function

For scalar functions $g(t)$ and $f(x, y)$ the **composite function** $g \circ f$ is defined by

$$(g \circ f)(x, y) = g(f(x, y))$$

for all $(x, y) \in D(f)$ for which $f(x, y) \in D(g)$.

When composing multivariable functions, it is very important to make sure the range of the inner function is a subset of the domain of the outer function. For example, we cannot compose scalar functions $f(x, y)$ and $h(x, y)$ since $f(x, y) \in \mathbb{R}$ which is not acceptable input into h.

We shall refer to the following theorems collectively as the **Continuity Theorems**.

THEOREM 1

If f and g are both continuous at (a, b), then $f + g$ and fg are continuous at (a, b).

Proof: We prove the result for $f + g$ and leave the proof for fg as an exercise. By the hypothesis and the definition of continuous function we have that

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

Hence, by definition of the sum and limit properties, we get

$$\begin{aligned}\lim_{(x,y)\to(a,b)} (f + g)(x, y) &= \lim_{(x,y)\to(a,b)} f(x, y) + \lim_{(x,y)\to(a,b)} g(x, y)\\ &= f(a, b) + g(a, b)\\ &= (f + g)(a, b)\end{aligned}$$

□

EXERCISE 1

Complete the proof of the theorem by proving that fg is continuous at (a, b).

THEOREM 2

If f and g are both continuous at (a, b) and $g(a, b) \neq 0$, then the quotient $\frac{f}{g}$ is continuous at (a, b).

EXERCISE 2

Use the Limit Theorems to prove Theorem 2. Where is the hypothesis $g(a, b) \neq 0$ used explicitly?

THEOREM 3 If $f(x, y)$ is continuous at (a, b) and $g(t)$ is continuous at $f(a, b)$, then the composition $g \circ f$ is continuous at (a, b).

Proof: Let $\epsilon > 0$. By definition of continuity we have that

$$\lim_{t \to f(a,b)} g(t) = g(f(a, b))$$

So, by definition of a limit there exists a $\delta_1 > 0$ such that

$$|t - f(a, b)| < \delta_1 \quad \text{implies} \quad |g(t) - g(f(a, b))| < \epsilon \tag{3.1}$$

Similarly, we have that

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

Hence, given the above δ_1, there exists a $\delta > 0$ such that

$$\|(x, y) - (a, b)\| < \delta \quad \text{implies} \quad |f(x, y) - f(a, b)| < \delta_1 \tag{3.2}$$

Notice that the conclusion of (3.2) is the hypothesis of (3.1) where $t = f(x, y)$. Hence, combining (3.1) and (3.2), we get

$$\|(x, y) - (a, b)\| < \delta \text{ implies } |f(x, y) - f(a, b)| < \delta_1 \text{ implies } |g(f(x, y)) - g(f(a, b))| < \epsilon$$

or equivalently,

$$\|(x, y) - (a, b)\| < \delta \quad \text{implies} \quad |(g \circ f)(x, y) - (g \circ f)(a, b)| < \epsilon$$

Consequently, by definition of a limit,

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

which proves that $g \circ f$ is continuous at (a, b). □

Before we can apply these theorems, we need a list of **basic functions** which are known to be continuous on their domains:

- the constant function $f(x, y) = k$
- the power functions $f(x, y) = x^n$, $f(x, y) = y^n$
- the logarithm function $\ln(\cdot)$
- the exponential function $e^{(\cdot)}$
- the trigonometric functions, $\sin(\cdot)$, $\cos(\cdot)$, etc.
- the inverse trigonometric functions, $\arcsin(\cdot)$, etc.
- the absolute value function $|\cdot|$

EXERCISE 3 Prove that the constant function $f(x, y) = k$ and the coordinate functions $f(x, y) = x$, $f(x, y) = y$ are continuous on their domains.

EXAMPLE 1

Prove that $h(x, y) = \sin(6x^2y + 3xy^2)$ is continuous for all $(x, y) \in \mathbb{R}^2$.

Solution: By applying Theorem 1 to the constant function and power functions, it follows that

$$f(x, y) = 6x^2y + 3xy^2 \tag{3.3}$$

is continuous for all $(x, y) \in \mathbb{R}^2$. Theorem 3, with $g(\cdot) = \sin(\cdot)$ and f as in equation (3.3), now implies that h is continuous for all $(x, y) \in \mathbb{R}^2$.

EXERCISE 4

Prove that $h(x, y) = (xy)^\pi$ is continuous for all (x, y) which satisfy $xy > 0$. Which of the theorems and basic functions do you have to use?

EXERCISE 5

Which of the basic functions and theorems do you have to use in order to prove that $h(x, y) = \dfrac{\sin^2 |\, x + 2y \,|}{x^2 + y^2}$ is continuous for all $(x, y) \neq (0, 0)$?

These examples and exercises show that by using the Continuity Theorems, one can often prove continuity of a given function essentially "by inspection". However, for certain points, where the Continuity Theorems can not be applied, one still has to use the definition of continuity in order to determine whether or not the function is continuous. Here is an example:

EXAMPLE 2

Discuss the continuity of the function f defined by

$$f(x, y) = \begin{cases} \dfrac{e^{xy} - 1}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{if } (x, y) = (0, 0) \end{cases}$$

Solution: For $(x, y) \neq (0, 0)$ the Continuity Theorems immediately imply that f is continuous at these points.

Observe the point $(0, 0)$ is singled out in the definition of the function. Thus, the Continuity Theorems cannot be applied at $(0, 0)$ and so we have to use the definition. That is, we have to determine whether

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

On the line $y = x$ we get

$$\lim_{(x,y)\to(0,0)} f(x, x) = \lim_{x\to 0} \frac{e^{x^2} - 1}{2x^2} = \lim_{x\to 0} \frac{2xe^{x^2}}{4x} = \lim_{x\to 0} \frac{e^{x^2}}{2} = \frac{1}{2}$$

by L'Hôpital's rule. It follows that $\lim\limits_{(x,y)\to(0,0)} f(x, y)$ does not equal $f(0, 0)$, and hence by definition, f is not continuous at $(0, 0)$.

EXERCISE 6 Would the function f in Example 2 be continuous at $(0,0)$ if we defined $f(0,0) = \frac{1}{2}$?

EXAMPLE 3 Discuss the continuity of the function f defined by

$$f(x,y) = \begin{cases} \dfrac{|y-x|}{y-x} & \text{if } x \neq y \\ 0 & \text{if } x = y \end{cases}$$

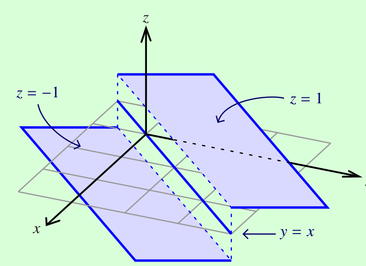
Solution: For points (x,y) with $x \neq y$ the Continuity Theorems immediately imply that f is continuous at these points.

We can not apply the continuity theorems at the points (x,y) with $x = y$. Consider any one of these points and denote it by (a,a).

If (x,y) approaches (a,a) with $y - x > 0$, then $|y - x| = y - x$, and $f(x,y)$ approaches (and in fact equals) 1. On the other hand, if (x,y) approaches (a,a) with $y - x < 0$, then $f(x,y)$ approaches -1. Thus,

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

does not exist. So, by definition of continuity, f is not continuous at (a,a).
The geometric interpretation is simple. The graph of f consists of two parallel half-planes which form a "step" along the line $y = x$.



3.3 Limits Revisited

So far in this chapter, we have shown how to prove that a function is continuous at a point essentially "by inspection" by using the Continuity Theorems. This makes it easy to evaluate $\lim\limits_{(x,y)\to(a,b)} f(x,y)$ if f is continuous at (a,b). In particular, if f is continuous at (a,b), then $\lim\limits_{(x,y)\to(a,b)} f(x,y)$ can be evaluated simply by evaluating $f(a,b)$.

EXAMPLE 1 Define $f(x,y) = \dfrac{\cos\sqrt{x^2+y^2}}{x^2+y^2}$, for $(x,y) \neq (0,0)$. Evaluate

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

Solution: By the Continuity Theorems, f is continuous for all $(x,y) \neq (0,0)$. Thus, by definition of continuity,

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

EXERCISE 1 Evaluate $\lim_{(x,y)\to(1,\pi)} \ln(1 + e^{\sin xy})$ justifying your method.

REMARK

In applying the Squeeze Theorem one has to prove that $\lim_{(x,y)\to(a,b)} B(x, y) = 0$. One hopes to be able to evaluate this limit by inspection, and so one tries to set up the inequality in the Squeeze Theorem so that $B(x, y)$ is continuous at (a, b).

Chapter 3 Problem Set

1. Let $f(x,y) = \begin{cases} \dfrac{x^2 - y^2}{|x| + |y|} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$
   (a) Evaluate $\lim_{(x,y)\to(2,1)} f(x, y)$.
   (b) Determine if $f(x, y)$ is continuous at $(0, 0)$.
2. Let $f(x,y) = \begin{cases} \dfrac{x^4 - y^4}{x^2 + y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$ Determine all points where f is continuous.
3. Let $f(x,y) = \begin{cases} \dfrac{x^2 - y^2}{x^2 + y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$ Determine all points where f is continuous.
4. For each function f, determine (with proof) whether or not $\lim_{(x,y)\to(0,0)} f(x, y)$ exists. Define $f(0, 0)$ so as to make the function continuous at $(0, 0)$, when possible.
   (a) $f(x,y) = \dfrac{x^3 - 2y^3}{x^2 + 2y^2}$
   (b) $f(x,y) = \dfrac{xy^4}{x^2 + y^6}$
   (c) $f(x,y) = \dfrac{xy^3}{x^2 + y^6}$
   (d) $f(x,y) = \dfrac{2|x| - |y|}{|x| + 2|y|}$
   (e) $f(x,y) = \dfrac{x^2 - 6y^2}{|x| + 3|y|}$
   (f) $f(x,y) = \dfrac{\sin(x^2 + 2y^2)}{x^2 + y^2}$
   (g) $f(x,y) = \dfrac{y^2 - 4|y| - 2|x|}{|x| + 2|y|}$.
5. Let $f(x,y) = \dfrac{\sin 2(x^2 + y^2)}{(x^2 + y^2)}$, for $(x, y) \neq (0, 0)$.
   (a) By using the inequality
   $$\theta - \frac{1}{6}\theta^3 \leq \sin\theta \leq \theta, \quad \text{for } \theta \geq 0$$
   or otherwise, evaluate $\lim_{(x,y)\to(0,0)} f(x, y)$.
   (b) Define $f(0, 0)$ so as to make the function f continuous at $(0, 0)$.
   (c) Use theorems on continuity to prove that the function f as defined in (b) is continuous for all $(x, y) \neq (0, 0)$.
6. A function f is defined by
   $$f(x,y) = \begin{cases} \dfrac{1 - e^{-|xy|}}{\sqrt{x^2 + y^2}} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$$
   Prove that f is continuous for all $(x, y) \in \mathbb{R}^2$.

   Hint: Don't do much work for $(x, y) \neq (0, 0)$. For $(x, y) = (0, 0)$, you may use the inequality $0 < 1 - e^{-u} < u$, for all $u > 0$.
7. Let $f(x,y) = \dfrac{y^2 - x^4}{y^2 + x^4}$, for $(x, y) \neq (0, 0)$.
   (a) Give the range of f, and sketch typical level curves $f(x, y) = k$. In your diagram, describe the set of points (x, y) for which $|f(x, y)| \leq \frac{3}{5}$.
   (b) On the basis of part (a), draw a conclusion about $\lim_{(x,y)\to(0,0)} f(x, y)$. Explain.

4. For each function f, determine (with proof) whether or not $\lim_{(x,y)\to(0,0)} f(x,y)$ exists. Define $f(0,0)$ so as to make the function continuous at $(0,0)$, when possible.

(a) $f(x,y) = \dfrac{x^3 - 2y^3}{x^2 + 2y^2}$

(b) $f(x,y) = \dfrac{xy^4}{x^2 + y^6}$

(c) $f(x,y) = \dfrac{xy^3}{x^2 + y^6}$

(d) $f(x,y) = \dfrac{2|x| - |y|}{|x| + 2|y|}$

(e) $f(x,y) = \dfrac{x^2 - 6y^2}{|x| + 3|y|}$

(f) $f(x,y) = \dfrac{\sin(x^2 + 2y^2)}{x^2 + y^2}$

(g) $f(x,y) = \dfrac{y^2 - 4|y| - 2|x|}{|x| + 2|y|.}$

(a) → approach along $y=0$

$\lim_{x\to0} x = 0$.

candidate limit : $L=0$

→ verify $|f(x,y)-L|$

$$\left|\frac{x^3-2y^3}{x^2+2y^2}\right| \le \frac{|x|^3+2|y|^3}{x^2+2y^2} = \frac{|x|^3}{x^2+2y^2} + \frac{2|y|^3}{x^2+2y^2}$$

$x^2+2y^2 \ge x^2$ $\quad \frac{x^2}{x^2+2y^2} \le 1$ $\quad \frac{|x|x^2}{x^2+2y^2} \le |x|$

Similarly $\frac{2|y|^3}{x^2+2y^2} \le 2|y|$

$|f(x,y)-L| \le |x|+2|y|$

$\lim_{(x,y)\to(0,0)} |x|+2|y| = 0$

∴ by squeeze theorem, $\lim f(x,y)$ exist.

(b) → approach along $y=mx$

$$\lim_{x\to0} \frac{m^4x^5}{x^2+m^6x^6} = \lim_{x\to0} \frac{x^2(m^4x^3)}{x^2(1+m^6x^4)} = \lim_{x\to0} \frac{m^4x^3}{1+m^6x^4} \overset{LHR}{=} 0.$$

candidate limit $L=0$

→ verify $|f(x,y)-L|=0$

$$\left|\frac{xy^4}{x^2+y^6}\right| = \frac{y^4|x|}{x^2+y^6}$$

$2|x||y^3| \le x^2+y^6$

$$\le \frac{|y|\cdot\frac12(x^2+y^6)}{x^2+y^6} = \frac12|y|$$

$\lim_{(x,y)\to(0,0)} \frac12|y| = 0$.

∴ by squeeze theo, lim exist

(c) → approach along $y=(mx)^{\frac13}$

$$\lim_{x\to0} \frac{x\cdot mx}{x^2+(mx)^2} = \lim_{x\to0} \frac{x^2\cdot m}{x^2(1+m^2)} = \frac{m}{1+m^2}$$

depend on m. DNE

(d) → approach along $y=0$

$\lim_{x\to0} \frac{2|x|}{|x|} = 2$

→ approach along $x=0$

$\lim_{y\to0} \frac{-|y|}{2|y|} = -\frac12$.

$2 \ne -\frac12$ DNE.

(e) → approach along $y=0$

$\lim_{x\to0} \frac{x^2}{|x|} = 0$

candidate limit : $L=0$

$$\left|\frac{x^2-6y^2}{|x|+3|y|}\right| \le \cancel{\frac{x^2+6xy+9y^2}{|x|+3|y|}}$$

法1 $$\left|\frac{x^2-6y^2}{|x|+3|y|}\right| = \left|\frac{x^2-9y^2+3y^2}{|x|+3|y|}\right| \le \frac{x^2-9y^2}{|x|+3|y|} + \frac{3y^2}{|x|+3|y|}$$

法2 : $$\left|\frac{x^2-6y^2}{|x|+3|y|}\right| \le \frac{x^2}{|x|+3|y|} - \frac{6y^2}{|x|+3|y|}$$

$$\lim_{(x,y)\to(0,0)} \frac{x^2}{|x|+3|y|} - \frac{6y^2}{|x|+3|y|}$$

$$= \lim_{(x,y)\to(0,0)} \frac{2x}{1} - \frac{12|y|}{3} \quad (\text{L'HR})$$

$= 0$.

∴ by squeeze theo. L exist

(f) $$\frac{-1}{x^2+y^2} \le \frac{\sin(x^2+2y^2)}{x^2+y^2} \le \frac{1}{x^2+y^2}$$

→ approach along $y=0$

$$\lim_{x\to0} \frac{\sin(x^2)}{x^2} = \lim_{x\to0} \frac{2x\cos(x^2)}{x^2} = \lim_{x\to0} \frac{-4\cancel{x^2}\sin(x^2)}{\cancel{x^2}} = -4$$

→ approach along $y=x$ $\qquad$ #

$$\lim_{x\to0} \frac{\sin(3x^2)}{x^2} = \lim_{x\to0} \frac{6\cancel{x}\cos(3x^2)}{2\cancel{x}} = 3\cos(3x^2) = 0.$$

(g) → approach along $y=0$

$\lim_{x\to0} \frac{-2|x|}{|x|} = -2$

$$\to |f(x,y)-L| = \lim_{(0,0)} \frac{y^2}{|x|+2|y|} - \frac{2\cancel{(|x|+2|y|)}}{\cancel{|x|+2|y|}} + 2$$

$$= \lim_{(0,0)} \frac{2|y|}{2} - 2 + 2$$

$= 0$

∴ by squeeze theo L exist

1. Let $f(x, y) = \begin{cases} \dfrac{x^2 - y^2}{|x| + |y|} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{if } (x, y) = (0, 0). \end{cases}$

   (a) Evaluate $\lim_{(x,y)\to(2,1)} f(x, y)$.

   (b) Determine if $f(x, y)$ is continuous at $(0, 0)$.

(a) $\lim_{x\to 2} \frac{x^2-1}{|x|+1} = \lim_{x\to 2} \frac{(|x|+1)(|x|-1)}{|x|+1}$

$= \lim_{x\to 2} |x|-1$

$= 1$

(b) → approach limit through $y=mx$

$\lim_{x\to 0} \frac{x^2-m^2x^2}{|x|+|mx|} = \lim_{x\to 0} \frac{|x|(|x|-m^2|x|)}{|x|(1+|m|)}$

$= \lim_{x\to 0} \frac{|x|\cancel{(1+|m|)}(1-|m|)}{\cancel{1+|m|}}$

$= 0$

→ let $f(x,y) = \frac{x^2-y^2}{|x|+|y|}$ $L=0$

$|f(x,y) - L| = \left|\frac{x^2-y^2}{|x|+|y|}\right|$

$= \left|\frac{(|x|+|y|)(|x|-|y|)}{|x|+|y|}\right|$

$= |x|+|y|$

→ $\lim_{(0,0)} |x|+|y| = 0$

∴ By squeeze theorem, $\lim = 0$

∴ cont.

2. Let $f(x, y) = \begin{cases} \dfrac{x^4 - y^4}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{if } (x, y) = (0, 0). \end{cases}$ Determine all points where f is continuous.

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

∴ $f(x,y)$ is cont. in all points

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

3. Let $f(x, y) = \begin{cases} \dfrac{x^2 - y^2}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{if } (x, y) = (0, 0). \end{cases}$ Determine all points where f is continuous.

→ approach limit through $y=mx$

$L_{00} \frac{x^2-y^2}{x^2+y^2} = L_{00} \frac{x^2-m^2x^2}{x^2+m^2x^2} = L_{00} \frac{x^2(1-m^2)}{x^2(1+m^2)}$

$= \frac{1-m^2}{1+m^2}$

limit DNE at $(0,0)$

→ cont at $(x,y) \neq (0,0)$

5. Let $f(x, y) = \dfrac{\sin 2(x^2 + y^2)}{(x^2 + y^2)}$, for $(x, y) \neq (0, 0)$.

   (a) By using the inequality

$$\theta - \frac{1}{6}\theta^3 \leq \sin\theta \leq \theta, \quad \text{for } \theta \geq 0$$

   or otherwise, evaluate $\lim_{(x,y)\to(0,0)} f(x, y)$.

   (b) Define $f(0, 0)$ so as to make the function f continuous at $(0, 0)$.

   (c) Use theorems on continuity to prove that the function f as defined in (b) is continuous for all $(x, y) \neq (0, 0)$.

(a) let $\theta = x^2+y^2$.

as $(x,y)\to(0,0)$, $\theta \to 0$.

$\lim_{(x,y)\to(0,0)} f(x,y) = \lim_{\theta\to 0} \frac{\sin 2\theta}{\theta}$

$2\theta - \frac{1}{6}(2\theta)^3 \leq \sin 2\theta \leq 2\theta$

$\frac{2\theta - \theta^3}{\theta} \leq \frac{\sin 2\theta}{\theta} \leq \frac{2\theta}{\theta}$

$2 - \theta^2 \leq \frac{\sin 2\theta}{\theta} \leq 2$

$\theta \to 0$. $L = 2$. by squeeze theo

(b) $f(0,0) = 2$.

(c) ∵ x^2+y^2 cont on $\mathbb{R}^2$

∴ $\frac{1}{x^2+y^2}$ cont on $(x,y) \neq (0,0)$

∵ $\sin(x)$ cont on $\mathbb{R}$.

∴ By continuous theorem 1&2

$\sin x$ cont $\forall (x,y) \neq (0,0)$

6. A function f is defined by

$$f(x,y) = \begin{cases} \dfrac{1-e^{-|xy|}}{\sqrt{x^2+y^2}} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$$

Prove that f is continuous for all $(x,y) \in \mathbb{R}^2$.

Hint: Don't do much work for $(x,y) \neq (0,0)$. For $(x,y) = (0,0)$, you may use the inequality $0 < 1 - e^{-u} < u$, for all $u > 0$.

$\rightarrow (x,y) \neq (0,0)$

cont. by cont. theorem 1 & 2

$\rightarrow (x,y) = (0,0)$

cosine inequality

$$\left| \frac{1-e^{-|xy|}}{\sqrt{x^2+y^2}} \right| < \frac{|xy|}{\sqrt{x^2+y^2}} \leq \frac{x^2+y^2}{\sqrt{x^2+y^2}} = \sqrt{x^2+y^2}$$

$$\lim_{(0,0)} \sqrt{x^2+y^2} = 0 = f(0,0)$$

$\therefore$ cont. at $(x,y) = (0,0)$

7. Let $f(x,y) = \dfrac{y^2 - x^4}{y^2 + x^4}$, for $(x,y) \neq (0,0)$.

(a) Give the range of f, and sketch typical level curves $f(x,y) = k$. In your diagram, describe the set of points (x,y) for which $|f(x,y)| \leq \frac{3}{5}$.

(b) On the basis of part (a), draw a conclusion about $\lim\limits_{(x,y)\to(0,0)} f(x,y)$. Explain.

(a) $\cancel{f(x,y) = \dfrac{(y+x^2)(y-x^2)}{y^2+x^4}}$

$\cancel{y^2 + x^2 > 0.}$

$\rightarrow x^4 \geq 0 \quad y^2 - x^4 \leq y^2 \leq y^2 + x^4.$

$$\frac{y^2-x^4}{y^2+x^4} \leq \frac{y^2+x^4}{y^2+x^4} = 1$$

$\rightarrow y^2 \geq 0 \quad x^4 \leq 0$

$y^2 - x^4 \leq y^2 + x^4$

$y^2 - x^4 \geq -y^2 - x^4$

$$\frac{y^2-x^4}{y^2+x^4} \geq -1$$

用contradiction 推导

$$\frac{y^2-x^4}{y^2+x^4} < -1$$

$y^2 - x^4 < -y^2 - x^4$

$2y^2 < 0 \quad \Rightarrow$ contradict

(b) DNE

Chapter 4

The Linear Approximation

4.1 Partial Derivatives

A scalar function $f(x, y)$ can be differentiated in two natural ways:

1. Treat y as a constant and differentiate with respect to x to obtain $\dfrac{\partial f}{\partial x}$.
2. Treat x as a constant and differentiate with respect to y to obtain $\dfrac{\partial f}{\partial y}$.

The derivatives $\dfrac{\partial f}{\partial x}$ and $\dfrac{\partial f}{\partial y}$ are called the (first) **partial derivatives** of f.

Here is the formal definition.

DEFINITION

Partial Derivatives

The **partial derivatives** of $f(x, y)$ are defined by

$$\frac{\partial f}{\partial x} = f_x = \lim_{h \to 0} \frac{f(x+h, y) - f(x, y)}{h}$$
$$\frac{\partial f}{\partial y} = f_y = \lim_{h \to 0} \frac{f(x, y+h) - f(x, y)}{h}$$

provided that these limits exist.

It is sometimes convenient to use **operator notation** $D_1 f$ and $D_2 f$ for the partial derivatives of $f(x, y)$. The notation $D_1 f$ means: differentiate f with respect to the variable in the first position, holding the other fixed. If the independent variables are x and y, then

$$D_1 f = \frac{\partial f}{\partial x} = f_x, \qquad D_2 f = \frac{\partial f}{\partial y} = f_y$$

Typically one tries to calculate the partial derivatives by using the standard rules for differentiation of functions of one variable. However, if these cannot be applied, then the definition of the partial derivatives must be used.

EXAMPLE 1 Consider the function f defined by $f(x,y) = xe^{kxy}$ where k is a constant. Determine $\dfrac{\partial f}{\partial x}$ and $\dfrac{\partial f}{\partial y}$.

Solution: By using the Product Rule and Chain Rule for differentiation,

$$\frac{\partial f}{\partial x} = (1)e^{kxy} + xe^{kxy}(ky) = (1+kxy)e^{kxy}$$
$$\frac{\partial f}{\partial y} = xe^{kxy}(kx) = kx^2e^{kxy}$$

EXERCISE 1 A function f is defined by $f(x,y) = \sin(xy^2)$. Determine f_x and f_y.

EXAMPLE 2 A function f is defined by $f(x,y) = (x^3+y^3)^{\frac{1}{3}}$. Determine whether $\dfrac{\partial f}{\partial x}(0,0)$ exists.

Solution: By single-variable differentiation rules,

$$\frac{\partial f}{\partial x}(x,y) = \frac{x^2}{(x^3+y^3)^{2/3}} \tag{4.1}$$

for all (x,y) such that $x^3+y^3 \neq 0$. One cannot substitute $(x,y) = (0,0)$ in equation (4.1) since the denominator would be zero. Thus, we must use the definition of the partial derivatives at $(0,0)$. We get

$$\begin{aligned}\frac{\partial f}{\partial x}(0,0) &= \lim_{h\to 0}\frac{f(0+h,0)-f(0,0)}{h}\\ &= \lim_{h\to 0}\frac{(h^3+0^3)^{1/3}-0}{h}\\ &= \lim_{h\to 0} 1 = 1\end{aligned}$$

EXAMPLE 3 Let $f(x,y) = \begin{cases}\dfrac{xy}{x^2+y^2} & \text{if } (x,y)\neq(0,0)\\ 0 & \text{if } (x,y)=(0,0)\end{cases}$. Calculate $f_x(0,0)$ and $f_y(0,0)$.

Solution: Since f changes definition at $(0,0)$, we must use the definition of the partial derivatives. We get

$$f_x(0,0) = \lim_{h\to 0}\frac{f(0+h,0)-f(0,0)}{h} = \lim_{h\to 0}\frac{\frac{h(0)}{h^2+0^2}-0}{h} = 0$$
$$f_y(0,0) = \lim_{h\to 0}\frac{f(0,0+h)-f(0,0)}{h} = \lim_{h\to 0}\frac{\frac{0(h)}{0^2+h^2}-0}{h} = 0$$

REMARK

In Example 3.1.2, we showed that $f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$ is not continuous at $(0,0)$, but we have just shown that its partial derivatives exist! This demonstrates that the concept of partial derivatives does not match our concept of differentiability for functions of one variable from Calculus 1. We will look at this more in the next chapter.

EXERCISE 2 Refer to the function in Example 2. Show that $\dfrac{\partial f}{\partial x}(a,-a)$ does not exist for $a \neq 0$.

EXERCISE 3 A function f is defined by $f(x,y) = |x(y-1)|$. Determine whether $\dfrac{\partial f}{\partial x}(0,0)$ and $\dfrac{\partial f}{\partial x}(0,1)$ exist.

Generalization

We can extend what we have done for scalar functions of two variables to scalar functions of n variables $f(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^n$. To take the partial derivative of f with respect to its i-th variable, we hold all the other variables constant and differentiate with respect to the i-th variable.

EXAMPLE 4 Let $f(x,y,z) = xy^2z^3$. Find f_x, f_y, and f_z.

Solution: We have

$$\begin{aligned} f_x(x,y,z) &= y^2z^3 \\ f_y(x,y,z) &= 2xyz^3 \\ f_z(x,y,z) &= 3xy^2z^2 \end{aligned}$$

EXERCISE 4 For $f(x,y,z)$, write the precise definition of $\dfrac{\partial f}{\partial x}$, $\dfrac{\partial f}{\partial y}$, and $\dfrac{\partial f}{\partial z}$.

4.2 Higher-Order Partial Derivatives

Second Partial Derivatives

Observe that the partial derivatives of a scalar function of two variables are both scalar functions of two variables. Therefore, we can take the partial derivatives of the partial derivatives of any scalar function.

In how many ways can one calculate a second partial derivative of $f(x, y)$? Since both of the partial derivatives of f have two partial derivatives, there are four possible second partial derivatives of f. They are:

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right), \quad \text{i.e. differentiate} \quad \frac{\partial f}{\partial x} \quad \text{with respect to} \quad x, \quad \text{with} \quad y \quad \text{fixed.}$$

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right), \quad \text{i.e. differentiate} \quad \frac{\partial f}{\partial x} \quad \text{with respect to} \quad y, \quad \text{with} \quad x \quad \text{fixed.}$$

Similarly

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right), \qquad \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)$$

It is often convenient to use the subscript notation or the operator notation:

$$\frac{\partial^2 f}{\partial x^2} = f_{xx} = D_1^2 f, \qquad \frac{\partial^2 f}{\partial y \partial x} = f_{xy} = D_2 D_1 f$$

$$\frac{\partial^2 f}{\partial x \partial y} = f_{yx} = D_1 D_2 f, \qquad \frac{\partial^2 f}{\partial y^2} = f_{yy} = D_2^2 f$$

The subscript notation suggests that one could write the second partial derivatives in a 2×2 matrix.

DEFINITION

Hessian Matrix

The **Hessian matrix** of $f(x, y)$, denoted by $Hf(x, y)$, is defined as

$$Hf(x, y) = \begin{bmatrix} f_{xx}(x, y) & f_{xy}(x, y) \\ f_{yx}(x, y) & f_{yy}(x, y) \end{bmatrix}$$

EXAMPLE 1

Let k be a constant. Find all the second partial derivatives of $f(x, y) = xe^{kxy}$.

Solution: We first calculate the first partial derivatives. We have

$$\frac{\partial f}{\partial x}(x, y) = e^{kxy} + kxye^{kxy}$$
$$\frac{\partial f}{\partial y}(x, y) = kx^2 e^{kxy}$$

Thus, we get

$$\frac{\partial^2 f}{\partial x^2}(x,y) = \frac{\partial}{\partial x}\left[e^{kxy} + kxye^{kxy}\right] = 2kye^{kxy} + k^2xy^2e^{kxy}$$
$$\frac{\partial^2 f}{\partial y\partial x}(x,y) = \frac{\partial}{\partial y}\left[e^{kxy} + kxye^{kxy}\right] = 2kxe^{kxy} + k^2x^2ye^{kxy}$$
$$\frac{\partial^2 f}{\partial x\partial y}(x,y) = \frac{\partial}{\partial x}\left[kx^2e^{kxy}\right] = 2kxe^{kxy} + k^2x^2ye^{kxy}$$
$$\frac{\partial^2 f}{\partial y^2}(x,y) = \frac{\partial}{\partial y}\left[kx^2e^{kxy}\right] = k^2x^3e^{kxy}$$

In the previous example, observe that

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

This is in fact a general property of partial derivatives, subject to a continuity requirement, as follows.

THEOREM 1 **(Clairaut's Theorem)**

If f_{xy} and f_{yx} are defined in some neighborhood of (a,b) and are both continuous at (a,b), then

$$f_{xy}(a,b) = f_{yx}(a,b)$$

The proof is rather technical, and is thus omitted. If the continuity hypothesis on f_{xy} and f_{yx} is dropped, the theorem is no longer true. See Exercise 3 below.

EXERCISE 1 Verify that $f(x,y) = \ln(x^2+y^2)$ satisfies

$$f_{xx} + f_{yy} = 0, \qquad \text{for } (x,y) \neq (0,0)$$

EXERCISE 2 Verify that $f(x,y) = x^y$ satisfies

$$f_{xy} = f_{yx}, \qquad \text{for } x > 0$$

EXERCISE 3 Let $f(x,y) = \begin{cases} xy\dfrac{x^2-y^2}{x^2+y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$

(a) Show that $f_x(0,y) = -y$ and $f_y(x,0) = x$.

(b) Show that $f_{xy}(0,0) = -1$ and $f_{yx}(0,0) = 1$. Hence $f_{xy} \neq f_{yx}$.

(c) Explain why this does not contradict Clairaut's theorem.

Higher-Order Partial Derivatives

Of course, we can take higher-order partial derivatives in the expected way. In particular, observe that $f(x,y)$ has eight third partial derivatives. They are

$$f_{xxx}, \quad f_{xxy}, \quad f_{xyx}, \quad f_{xyy}, \quad f_{yxx}, \quad f_{yxy}, \quad f_{yyx}, \quad f_{yyy}$$

Not surprisingly, Clairaut's Theorem extends to higher-order partial derivatives as well. That is, if they are defined in a neighborhood of a point (a,b) and are continuous at (a,b), then the higher-order partial derivatives are equal regardless of the order the partial derivatives are taken. For example,

$$f_{xxy}(a,b) = f_{xyx}(a,b) = f_{yxx}(a,b)$$

For many situations, we will want to require that a function have continuous partial derivatives of some order. Thus, we introduce some notation for this.

If the k-th partial derivatives of $f(x_1,\ldots,x_n)$ are continuous, then we write

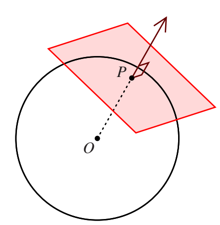
$$f \in C^k$$

and say "f is in class C^k."

So, $f(x,y) \in C^2$ means that f has continuous second partial derivatives, and therefore, by Clairaut's Theorem, we have that $f_{xy} = f_{yx}$.

4.3 The Tangent Plane

The surface of a sphere has a tangent plane at each point P, namely the plane through P that is orthogonal to the line joining P and the centre O. The tangent plane at P can be thought of as the plane which best approximates the surface of the sphere near P.



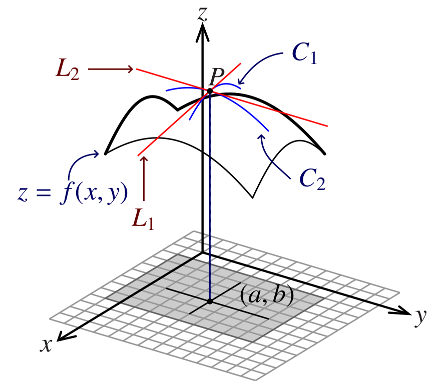
This concept can be generalized to a surface defined by an equation of the form

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

Let C_1 be the cross section $y = b$ of the surface, that is, C_1 is given by

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

It follows that $\dfrac{\partial f}{\partial x}(a,b)$ equals the slope of the tangent line L_1 of C_1 at the point $P(a,b,f(a,b))$. A similar interpretation holds for $\dfrac{\partial f}{\partial y}(a,b)$ in terms of the cross section $z=f(a,y)$.
We provisionally define the tangent plane to the surface $z = f(x,y)$ at the point $P(a,b,f(a,b))$ to be the plane which contains the tangent lines L_1 and L_2 (refer to the figure).



In order to derive the equation of the tangent plane, we note that any (non-vertical) plane through the point $P(a,b,f(a,b))$ has an equation of the form

$$z = f(a,b) + m(x-a) + n(y-b)$$

where m and n are constants. The intercept of this plane with the vertical plane $y=b$ is the line

$$z = f(a,b) + m(x-a) \tag{4.2}$$

We require this line to coincide with L_1. Thus the slope m of the line (4.2) must equal the slope $\dfrac{\partial f}{\partial x}(a,b)$ of the line L_1:

$$m = \frac{\partial f}{\partial x}(a,b)$$

A similar argument yields

$$n = \frac{\partial f}{\partial y}(a,b)$$

We make the following definition which we will formalize in Chapter 5.

DEFINITION
Tangent Plane

The **tangent plane** to $z = f(x,y)$ at the point $(a,b,f(a,b))$ is

$$z = f(a,b) + \frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b)$$

EXERCISE 1

The graph of the function

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

is the cone $z = \sqrt{x^2+y^2}$. Find the equation of the tangent plane at the point $(3,-4,5)$.

EXERCISE 2

Show that the tangent plane at any point on the cone in Exercise 1 passes through the origin.

REMARK

In Exercise 2, you should note that a tangent plane does not exist at the vertex $(0, 0, 0)$ of the cone, since the cone is not "smooth" there. We shall discuss the question of the existence of a tangent plane in Chapter 5.

4.4 Linear Approximation for $z = f(x, y)$

Review of the 1-D case

For a function $f(x)$ the tangent line can be used to approximate the graph of the function near the point of tangency. Recall that the equation of the tangent line to $y = f(x)$ at the point $(a, f(a))$ is

$$y = f(a) + f'(a)(x - a)$$

The function L_a defined by

$$L_a(x) = f(a) + f'(a)(x - a)$$

is called the **linearization** of f at a since $L_a(x)$ approximates $f(x)$ for x sufficiently close to a.

For x sufficiently close to a, the approximation

$$f(x) \approx L_a(x)$$

is called the **linear approximation** of f at a.

EXERCISE 1

Verify each approximation:

(a) $\sin x \approx x$, for x sufficiently close to 0,
(b) $\sqrt{1 + x} \approx 1 + \frac{1}{2}x$, for x sufficiently close to 0,
(c) $\ln x \approx (x - 1)$, for x sufficiently close to 1.

The 2-D case

For a function $f(x, y)$, the tangent plane can be used to approximate the surface $z = f(x, y)$ near the point of tangency.

DEFINITION

Linearization

Linear Approximation

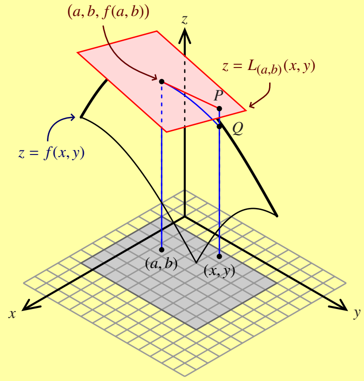
For a function $f(x,y)$ we define the **linearization** $L_{(a,b)}(x,y)$ of f at (a,b) by

$$L_{(a,b)}(x,y) = f(a,b) + \frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b)$$

We call the approximation

$$f(x,y) \approx L_{(a,b)}(x,y)$$

the **linear approximation** of $f(x,y)$ at (a,b).



EXAMPLE 1 Use the linear approximation to approximate $\sqrt{(0.95)^3 + (1.98)^3}$.

Solution: A choice of function and point of tangency must be made. Let

$$f(x,y) = \sqrt{x^3 + y^3}, \qquad \text{and} \qquad (a,b) = (1,2)$$

The partial derivatives of f are

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

Thus, the linear approximation is

$$\begin{aligned} f(x,y) &\approx L_{(1,2)}(x,y) \\ &= f(1,2) + f_x(1,2)(x-1) + f_y(1,2)(y-2) \\ &= 3 + \frac{1}{2}(x-1) + 2(y-2) \end{aligned} \tag{4.3}$$

Hence,

$$\sqrt{(0.95)^3 + (1.98)^3} = f(0.95, 1.98) \approx 3 + \frac{1}{2}(-0.05) + 2(-0.02) = 2.935$$

Note: The calculator value is 2.935943.

REMARK

Resist the temptation to expand the brackets and simplify in equation (4.3). The bracketed terms represent small increments, and it is helpful to keep them separate.

EXERCISE 2 Calculate $\sqrt{\sin\left(\frac{1}{10}\right) + \tan\left(\frac{3}{4}\right)}$ approximately. Compare your answer with the value from a calculator.

Hint: Choose the point of tangency so that the increments in x and y do not exceed $\frac{1}{10}$. Use the approximate value 3.14 for π.

EXERCISE 3

Verify each approximation:

(a) $\dfrac{xy}{x+y} \approx \dfrac{6}{5} + \dfrac{9}{25}(x-2) + \dfrac{4}{25}(y-3)$, for (x, y) sufficiently close to $(2, 3)$

(b) $\ln(x^2 + y) \approx 2(x-1) + y$, for (x, y) sufficiently close to $(1, 0)$

(c) $e^{3x-2y} \approx 1 + 3x - 2y$, for (x, y) sufficiently close to $(0, 0)$.

Increment Form of the Linear Approximation

Suppose that we know $f(a, b)$ and want to calculate $f(x, y)$ at a nearby point. Let

$$\Delta x = x - a, \quad \Delta y = y - b$$

and

$$\Delta f = f(x, y) - f(a, b)$$

The linear approximation is

$$f(x, y) \approx f(a, b) + \frac{\partial f}{\partial x}(a, b)(x - a) + \frac{\partial f}{\partial y}(a, b)(y - b)$$

This can be rearranged to yield

$$\Delta f \approx \frac{\partial f}{\partial x}(a, b)\Delta x + \frac{\partial f}{\partial y}(a, b)\Delta y \tag{4.4}$$

This gives an approximation for the change Δf in $f(x, y)$ due to a change $(\Delta x, \Delta y)$ away from the point (a, b).

We shall refer to equation (4.4) as the **increment form** of the linear approximation.

EXERCISE 4

An isosceles triangle has base 4 m, and equal angles of $\frac{\pi}{4}$. If the base is increased by 16 cm, and the equal angles are decreased by 0.1 radians, estimate the change in area.

4.5 Linear Approximation in Higher Dimensions

Linear Approximation in $\mathbb{R}^3$

Consider a function $f(x, y, z)$. By analogy with the case of a function of two variables, we define the linearization of f at $\mathbf{a} = (a, b, c)$ by

$$L_{\mathbf{a}}(x, y, z) = f(\mathbf{a}) + f_x(\mathbf{a})(x - a) + f_y(\mathbf{a})(y - b) + f_z(\mathbf{a})(z - c)$$

The notation is becoming cumbersome, but one can improve matters by noting that the final three terms can be represented by the dot product of the vectors

$$(x-a, y-b, z-c) = (x,y,z) - (a,b,c), \quad \text{and} \quad \nabla f(\mathbf{a}) = \big(f_x(\mathbf{a}), f_y(\mathbf{a}), f_z(\mathbf{a})\big)$$

The second vector is called the **gradient** of f at $\mathbf{a}$.

Here are the formal definitions.

DEFINITION

Gradient

Suppose that $f(x,y,z)$ has partial derivatives at $\mathbf{a} \in \mathbb{R}^3$. The **gradient** of f at $\mathbf{a}$ is defined by

$$\nabla f(\mathbf{a}) = \big(f_x(\mathbf{a}), f_y(\mathbf{a}), f_z(\mathbf{a})\big)$$

DEFINITION

Linearization

Linear Approximation

Suppose that $f(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^3$, has partial derivatives at $\mathbf{a} \in \mathbb{R}^3$. The **linearization** of f at $\mathbf{a}$ is defined by

$$L_{\mathbf{a}}(\mathbf{x}) = f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a}) \tag{4.5}$$

The **linear approximation** of f at $\mathbf{a}$ is

$$f(\mathbf{x}) \approx f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a}) \tag{4.6}$$

EXAMPLE 1

Consider the function f defined by

$$f(x,y,z) = \sqrt{x^2+y^2+z^2}$$

Find the gradient of f and the linear approximation for f at $\mathbf{a} = (1,2,-2)$.

Solution: Differentiate to obtain

$$\nabla f(x,y,z) = \left(\frac{x}{\sqrt{x^2+y^2+z^2}}, \frac{y}{\sqrt{x^2+y^2+z^2}}, \frac{z}{\sqrt{x^2+y^2+z^2}}\right)$$

Now, evaluate $\nabla f(x,y,z)$ at $\mathbf{a} = (1,2,-2)$ to get

$$\nabla f(\mathbf{a}) = \left(\frac{1}{3}, \frac{2}{3}, -\frac{2}{3}\right)$$

Thus,

$$\begin{aligned} L_{\mathbf{a}}(\mathbf{x}) &= f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a}) \\ &= 3 + \frac{1}{3}(x-1) + \frac{2}{3}(y-2) - \frac{2}{3}(z+2) \end{aligned}$$

So, the linear approximation for f at $(1,2,-2)$ is

$$f(x,y,z) \approx 3 + \frac{1}{3}(x-1) + \frac{2}{3}(y-2) - \frac{2}{3}(z+2)$$

EXERCISE 1 Use the linear approximation to estimate $4.99 \times 7.01 \times 9.99$. Compare your answer to the calculator value.

Linear Approximation in $\mathbb{R}^n$

The advantage of using vector notation is that equations (4.5) and (4.6) hold for a function of n variables $f(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^n$. For arbitrary $\mathbf{a} \in \mathbb{R}^n$, we have

$$\mathbf{x} - \mathbf{a} = (x_1 - a_1, x_2 - a_2, \ldots, x_n - a_n)$$

and we define the gradient of f at $\mathbf{a}$ to be

$$\nabla f(\mathbf{a}) = (D_1 f(\mathbf{a}), D_2 f(\mathbf{a}), \cdots, D_n f(\mathbf{a}))$$

Then, the increment form of the linear approximation for $f(\mathbf{x})$ is

$$\Delta f \approx \nabla f(\mathbf{a}) \cdot \Delta\mathbf{x}$$

Observe that this formula even works when $n = 1$. That is, for a function $g(t)$ of one variable this gives $\nabla g(a) = g'(a)$ and the increment form of the linear approximation is

$$\Delta g \approx \nabla g(a) \cdot \Delta x = g'(a)(x - a)$$

which is our familiar formula from Calculus 1.

For $f(x, y)$ we have $\nabla f(a, b) = (f_x(a, b), f_y(a, b))$ and the increment form of the linear approximation is

$$\Delta f \approx \nabla f(a, b) \cdot \Delta(x, y) = f_x(a, b)(x - a) + f_y(a, b)(y - b)$$

which matches our work above. Hence, we see that this is a true generalization.

Chapter 4 Problem Set

1. Let $f(x, y) = \begin{cases} \dfrac{\sin(xy)}{\ln(x^2 + y^2 + 1)} & \text{for } (x, y) \neq (0, 0) \\ 0 & \text{for } (x, y) = (0, 0). \end{cases}$
   (a) Prove that $f_x(0, 0)$ and $f_y(0, 0)$ exist.
   (b) Prove that f is not continuous at $(0, 0)$.

2. Let $f(x, y) = \begin{cases} \dfrac{x^3}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{if } (x, y) = (0, 0). \end{cases}$
   (a) Determine all points where $f(x, y)$ is continuous.
   (b) Find $f_x(0, 0)$ and $f_y(0, 0)$.

3. Find a function $g(x, y)$ such that $g(x, y)$ is continuous at $(0, 0)$, but $g_x(0, 0)$ and $g_y(0, 0)$ both do not exist. Justify your answer.

4. Find $f_x(0, 0)$ and $f_y(0, 0)$ for
$$f(x, y) = \begin{cases} \dfrac{x^3 - y^3}{x^2 + y^2} + 1 & \text{if } (x, y) \neq (0, 0) \\ 1 & \text{if } (x, y) = (0, 0) \end{cases}$$

5. Let $f(x, y) = \dfrac{xy}{x^2 + y^2}$.
   (a) Find the equation of the tangent plane of f at $(1, 2, 2/5)$.
   (b) Approximate $f(0.9, 2.1)$.

6. Find the linearization of the function at the given point.

(a) $f(x,y) = \ln(x+2y), \quad (a,b) = (3,-1)$
(b) $f(x,y) = \sqrt{\sin 3x + 4\tan y}, \quad (a,b) = (0, \frac{\pi}{4})$
(c) $f(x,y,z) = e^{x+2y+3z}, \quad (a,b,c) = (1,1,-1)$
(d) $f(x,y,z) = \ln(x^2 - yz), \quad (a,b,c) = (2,1,3)$

7. Use the linear approximation to approximate:
(a) $(0.99e^{0.02})^8$
(b) $\sqrt{(4.02)^2 + (3.95)^2 + (2.01)^2}$
(c) $\sqrt{e^{0.1} + 3\sin(0.05)}$

Compare your answers with the value from a calculator.

8. Find the first and second partial derivatives of
(a) $f(x,y) = \sqrt{2x^2 - y}$
(b) $g(x,y) = xe^{x+\cos y}$

9. Let $f(x,y) = \dfrac{xy}{x^2+y^2}$.
(a) Find the equation of the tangent plane of f at $(2,1,2/5)$.
(b) Approximate $f(1.9, 1.1)$.

10. A function g is defined by $g(x,t) = f(x-3t)$ where f is a function of one variable. If $f'(2) = 3$, calculate $g_x(5,1)$, $g_t(5,1)$. Show that $g_t(x,t) = -3g_x(x,t)$ in general.

11. A function f is defined by $f(x,y) = ye^{\frac{x}{y}}$, $y \neq 0$. Verify that the second mixed partial derivatives are equal:

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

12. Determine the values of the constants α and β for which the function $u(x,t) = e^{\alpha t}\sin\beta x$ satisfies the 1-d heat equation

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

13. (a) Consider f defined by

$$f(x,y) = \begin{cases} x\ln(x^2+2y^2), & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$$

Prove that f_y is defined for all $(x,y) \in \mathbb{R}^2$, but that f_y is not continuous at $(0,0)$.
(b) Invent another function with this property.

14. Let $f(x,y) = \sqrt{|xy|}$.
(a) Calculate $\frac{\partial f}{\partial x}(1,-4)$, $\frac{\partial f}{\partial x}(0,0)$, $\frac{\partial f}{\partial x}(0,1)$ if they exist. At which of these points is it necessary to use the definition of the partial derivative?
(b) At what points do the partial derivatives of f not exist? Make a conjecture based on part (a), and give a proof.

15. The temperature of a metal rod at position x, $0 \le x \le 1$, and at time t, $t \ge 0$ is given by $u(t,x) = 100e^{-t}\sin\pi x$. Find the rate of change of temperature with respect to position when $x = \frac{3}{4}$ and $t = 1$. Find the rate of change of temperature with respect to time when $x = \frac{3}{4}$ and $t = 1$. Illustrate these rates of change by sketching the cross sections $x = \frac{3}{4}$ and $t = 1$.

16. Let $u(x,t)$ denote the displacement (in mm) of a vibrating string at a point x on the string at time t. How would you physically interpret the functions $u_t(x,t)$ and $u_x(x,t)$?

17. A silo consists of a circular cylinder of radius 5 meters, and height 25 meters, capped by a hemisphere. Suppose that the radius is decreased by 5 centimeters and the height of the cylinder is increased by 10 centimeters. Use the linear approximation to estimate the change in volume.

18. If three resistors R_1, R_2, R_3 are connected in parallel, the total electrical resistance R is determined by

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}$$

If R_1, R_2 and R_3 initially equal 100, 200 and 300 ohms, and are increased by 1,2,4 ohms respectively, use the linear approximation to calculate the change in R. Compare with a direct calculation on a calculator.

19. Find all planes which are tangent to the surface $z = 1 - x^2 - y^2$, and contain the line passing through the points $(1,0,2)$ and $(0,2,2)$.

20. (a) Verify that $u = Ae^{-(x-ct)^2}$, where A and c are constants, satisfies the 1-d wave equation

$$u_{tt} = c^2 u_{xx} \qquad (*)$$

(b) Graph u versus x for $t = 0, \frac{1}{c}, \frac{2}{c}, \frac{3}{c}$, on the same axes. With what speed does the wave move along the x-axis?
(c) Find a solution of $(*)$ which describes a wave moving to the *left* along the x-axis.
(d) Let f be a function of one variable with a continuous second derivative. Verify that $u = f(x-ct)$ is a solution of the wave equation $(*)$.

21. Show that $u(x,t) = \int_0^{\frac{x}{2\sqrt{t}}} e^{-s^2}\,ds$ satisfies the 1-d heat equation $\dfrac{\partial u}{\partial t} = \dfrac{\partial^2 u}{\partial x^2}$. Sketch the level curves of $u(x,t)$.

22. * (a) Give a function $f(x,y)$ such that $f_{yx} = 0$.
(b) Find all functions $f(x,y)$ which have continuous second partial derivatives, and satisfy $f_{yx} = 0$.
(c) Suppose that $u(x,t)$ is a function which has continuous second partial derivatives on $\mathbb{R}^2$ and which satisfies the one dimensional wave equation

$$u_{tt} = c^2 u_{xx} \qquad (*)$$

where c is a constant. Determine how equation $(*)$ is transformed under the change of independent variables expressed by $p = x + ct$, $q = x - ct$. Using your answer to part (b), obtain the general solution of the wave equation $(*)$, and compare your answer with the special solutions discussed in # 20.

1. Let $f(x,y) = \begin{cases} \dfrac{\sin(xy)}{\ln(x^2+y^2+1)} & \text{for } (x,y) \neq (0,0) \\ 0 & \text{for } (x,y) = (0,0). \end{cases}$

(a) Prove that $f_x(0,0)$ and $f_y(0,0)$ exist.

(b) Prove that f is not continuous at $(0,0)$.

show $\lim_{(x,y)\to(0,0)} f(x,y)$ DNE or $\neq 0$.

(a) $f_x = \lim_{h\to 0} \dfrac{f(x+h,y)-f(x,y)}{h}$

$= \lim_{h\to 0} \dfrac{f(0+h,y)-f(0,0)}{h}$

$= \lim_{h\to 0} \dfrac{0}{h}$

$=0$

f_y 同理

(b) approach limit through $y=x$

$\lim_{(x,y)\to(0,0)} \dfrac{\sin(xy)}{\ln(x^2+y^2+1)}$

$= \lim_{x\to 0} \dfrac{\sin(x^2)}{\ln(2x^2+1)}$

$= \lim_{x\to 0} \cancel{2x}\cdot\cos(x^2)\cdot\dfrac{2x^2+1}{\cancel{4x}\ 2}$ (LHR)

$= \lim_{x\to 0} \dfrac{(2x^2+1)\cos(x^2)}{2}$

$= \dfrac{1}{2}$

$\because f(0,0)=0 \therefore f$ is not cont at $(0,0)$

2. Let $f(x,y) = \begin{cases} \dfrac{x^3}{x^2+y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$

(a) Determine all points where $f(x,y)$ is continuous.

(b) Find $f_x(0,0)$ and $f_y(0,0)$.

(a) $\dfrac{x^3}{x^2+y^2}$ is cont $\forall (x,y)\neq(0,0)$

check whether $f(x,y)$ cont at $(0,0)$

$\rightarrow$ let $f(x,y)=\dfrac{x^3}{x^2+y^2}$. $L=0$.

$|f(x,y)-L| = \left|\dfrac{x^3}{x^2+y^2}\right| \leq \left|\dfrac{x^3}{x^2}\right| = |x|$

$(\because y^2 \geq 0)$

let $B(x,y)=|x|$

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

$\therefore 0 \leq |f(x,y)-L| \leq 0$

by squeeze theorem $\lim_{(0,0)} f(x,y)=0$

$\rightarrow$ continuous. everywhere

(b) $f_x = \dfrac{3x^2(x^2+y^2)-x^3\cdot 2x}{(x^2+y^2)^2}$

$= \dfrac{3x^4+3x^2y^2-2x^4}{(x^2+y^2)^2}$

$= \dfrac{x^4}{x^4}$ (as $y\to 0$)

$=1$

$f_y = \dfrac{x^3\cdot 2y}{(x^2+y^2)^2}$

$= \dfrac{0}{y^4}$ (as $x\to 0$)

$=0$

3. Find a function $g(x,y)$ such that $g(x,y)$ is continuous at $(0,0)$, but $g_x(0,0)$ and $g_y(0,0)$ both do not exist. Justify your answer.

$g(x,y) = |x|+|y|$

V

4. Find $f_x(0,0)$ and $f_y(0,0)$ for

$$f(x,y) = \begin{cases} \dfrac{x^3 - y^3}{x^2 + y^2} + 1 & \text{if } (x,y) \neq (0,0) \\ 1 & \text{if } (x,y) = (0,0) \end{cases}$$

$$f_x = \lim_{h \to 0} \frac{f(x+h,y) - f(x,y)}{h}$$
$$= \lim_{h \to 0} \frac{f(h,0) - f(0,0)}{h}$$
$$= \lim_{h \to 0} \frac{h+1-1}{h}$$
$$= 1$$

$$f_y = \lim_{h \to 0} \frac{f(x,y+h) - f(x,y)}{h}$$
$$= \lim_{h \to 0} \frac{\frac{-(y+h)^3}{(y+h)^2} + 1 - 1}{h}$$
$$= -1$$

5. Let $f(x,y) = \dfrac{xy}{x^2 + y^2}$.

(a) Find the equation of the tangent plane of f at $(1, 2, 2/5)$.

(b) Approximate $f(0.9, 2.1)$.

(a) $f(1,2) = \frac{2}{5}$

$$f_x(1,2) = \frac{y(x^2+y^2) - xy \cdot 2x}{(x^2+y^2)^2} = \frac{6}{25}$$

$$f_y(1,2) = \frac{x(x^2+y^2) - xy \cdot 2y}{(x^2+y^2)^2} = -\frac{3}{25}$$

$$z = \frac{2}{5} + \frac{6}{25}(x-1) - \frac{3}{25}(y-2)$$

(b) $x = 0.9$ $y = 2.1$

$$z = \frac{2}{5} + \frac{6}{25}(-0.1) - \frac{3}{25} \cdot (0.1) = 0.364.$$

7. Use the linear approximation to approximate:

(a) $(0.99e^{0.02})^8$

(b) $\sqrt{(4.02)^2 + (3.95)^2 + (2.01)^2}$

(c) $\sqrt{e^{0.1} + 3\sin(0.05)}$

Compare your answers with the value from a calculator.

(a) $f(x,y) = (xe^y)^8 = x^8 e^{8y}$

$f(1,0) = 1$

$f_x(1,0) = 8x^7 e^{8y} = 8$

$f_y(1,0) = x^8 \cdot 8 \cdot e^{8y} = 8$

$z = 1 + 8(x-1) + 8y = 8x + 8y - 7$

$x = 0.99$ $y = 0.02$ $z = 1.08$

(b) $f(x,y,z) = \sqrt{x^2+y^2+z^2}$

$f(4,4,2) = 6$

$f_x(4,4,2) = \frac{1}{2}(x^2+y^2+z^2)^{-\frac{1}{2}} \cdot 2x = \frac{2}{3}$

$f_y(4,4,2) = \frac{1}{2}(x^2+y^2+z^2)^{-\frac{1}{2}} \cdot 2y = \frac{2}{3}$

$f_z(4,4,2) = \frac{1}{2}(x^2+y^2+z^2)^{-\frac{1}{2}} \cdot 2z = \frac{1}{3}$

$g = 6 + \frac{2}{3}(x-4) + \frac{2}{3}(y-4) + \frac{1}{3}(z-2)$

$= \frac{2}{3}x + \frac{2}{3}y + \frac{1}{3}z$

$x = 4.02$ $y = 3.95$ $z = 2.01$ $g = 5.98$

(c) 同理

8. Find the first and second partial derivatives of

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

(b) $g(x,y) = xe^{x+\cos y}$

(a) • first

$f_x = \frac{1}{2}(2x^2-y)^{-\frac{1}{2}} \cdot 4x$

$= 2x(2x^2-y)^{-\frac{1}{2}}$

$f_y = \frac{1}{2}(2x^2-y)^{-\frac{1}{2}} \cdot (-1)$

$= -\frac{1}{2}(2x^2-y)^{-\frac{1}{2}}$

• second

$f_{xx} = 2(2x^2-y)^{-\frac{1}{2}} + 2x \cdot (-\frac{1}{2})(2x^2-y)^{-\frac{3}{2}} 4x$

$= 2(2x^2-y)^{-\frac{1}{2}} - 4x^2 (2x^2-y)^{-\frac{3}{2}}$

$f_{yx} = f_{xy} = 2x \cdot (-\frac{1}{2})(-1) \cdot (2x^2-y)^{-\frac{3}{2}}$

$= x(2x^2-y)^{-\frac{3}{2}}$

$f_{yy} = -\frac{1}{2} \cdot (-\frac{1}{2})(2x^2-y)^{-\frac{3}{2}}(-1)$

$= -\frac{1}{4}(2x^2-y)^{-\frac{3}{2}}$

(b) • first

$g_x = e^{x+\cos y} + x \cdot e^{(x+\cos y)}$

$g_y = -(x \sin y) e^{(x+\cos y)}$

• second

$g_{xx} = e^{x+\cos y} + e^{x+\cos y} + xe^{x+\cos y}$

$= (2+x)e^{x+\cos y}$

$g_{xy} = g_{yx} = -\sin y \cdot e^{(x+\cos y)} - x\sin y \cdot e^{(x+\cos y)}$

$= -(1+x)\sin y \cdot e^{(x+\cos y)}$

$g_{yy} = x\sin^2 y e^{x+\cos y} - x\cos y e^{x+\cos y}$

9. Let $f(x,y) = \dfrac{xy}{x^2+y^2}$.

(a) Find the equation of the tangent plane of f at $(2, 1, 2/5)$.

(b) Approximate $f(1.9, 1.1)$.

与5相同

10. A function g is defined by $g(x,t) = f(x-3t)$ where f is a function of one variable. If $f'(2) = 3$, calculate $g_x(5,1)$, $g_t(5,1)$. Show that $g_t(x,t) = -3g_x(x,t)$ in general.

综合系

~~$g_x(x,t) = \lim_{h\to 0} \frac{g(x+h,t) - g(x,t)}{h}$~~

~~$= \lim_{h\to 0} \frac{f(x+h-3t) - f(x-3t)}{h}$~~

~~$= \lim_{h\to 0} \frac{f(2+h) - f(2)}{h}$~~

$g_x(x,t) = f_x(x-3t) = f_x(x-3t) = f'(2) = 3$

$g_t(x,t) = f_t(x-3t) = -3 \cdot f'(x-3t) = -3 \cdot 3 = -9$

$= -3g_x(x-3t)$

$g_x(5,1) = f_x(5,3) = f'(2) = 3$

$g_t(5,1) = -3g_x(5,1) = -6$

11. A function f is defined by $f(x,y) = ye^{\frac{x}{y}}$, $y \neq 0$. Verify that the second mixed partial derivatives are equal:

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

xy^{-1}

$\text{LHS} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial}{\partial x}\left(\cancel{e^{\frac{x}{y}}} + y \cdot (-xy^{-2}) e^{\frac{x}{y}}\right)$

$= \frac{\partial}{\partial x}\left(\cancel{e^{\frac{x}{y}}} - \frac{x}{y} e^{\frac{x}{y}}\right)$

$= \cancel{\frac{1}{y} e^{\frac{x}{y}}} - \frac{x}{y^2} e^{\frac{x}{y}}.$

$\text{RHS} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial}{\partial y}\left(y \cdot \frac{1}{y} e^{\frac{x}{y}}\right)$

$= \frac{-xe^{\frac{x}{y}}}{y^2}$

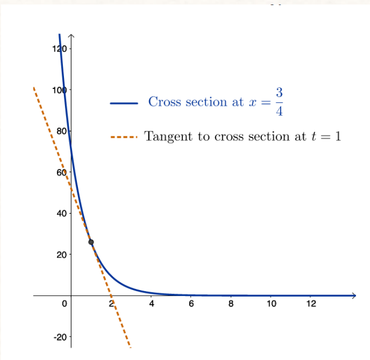
15. The temperature of a metal rod at position x, $0 \le x \le 1$, and at time t, $t \ge 0$ is given by $u(t, x) = 100e^{-t} \sin \pi x$. Find the rate of change of temperature with respect to position when $x = \frac{3}{4}$ and $t = 1$. Find the rate of change of temperature with respect to time when $x = \frac{3}{4}$ and $t = 1$. Illustrate these rates of change by sketching the cross sections $x = \frac{3}{4}$ and $t = 1$.

$u_x(t, x) = 100 e^{-t} \pi \cos(x \cdot \pi)$

$u_x(\frac{3}{4}, 1) = -81.72$

$u_t(t, x) = -100 e^{-t} \sin(\pi \cdot x)$

$u_x(1, \frac{3}{4}) = -26.01$



16. Let $u(x, t)$ denote the displacement (in mm) of a vibrating string at a point x on the string at time t. How would you physically interpret the functions $u_t(x, t)$ and $u_x(x, t)$?

amplitude

Hz.

用物理意义

17. A silo consists of a circular cylinder of radius 5 meters, and height 25 meters, capped by a hemisphere. Suppose that the radius is decreased by 5 centimeters and the height of the cylinder is increased by 10 centimeters. Use the linear approximation to estimate the change in volume.

$V(r, h) = \pi r^2 h$.

$V(5, 25) = 625\pi$

$V_r(5, 25) = 2\pi r h = 250\pi$

$V_h(5, 25) = \pi r^2 = 25\pi$

$\Delta V = 250\pi \cdot \Delta r + 25\pi \Delta h = -250\pi \cdot 0.05 + 25\pi \cdot 0.1 = \cancel{-10\pi}$

答案是-39.3

18. If three resistors R_1, R_2, R_3 are connected in parallel, the total electrical resistance R is determined by

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}$$

If R_1, R_2 and R_3 initially equal 100, 200 and 300 ohms, and are increased by 1,2,4 ohms respectively, use the linear approximation to calculate the change in R. Compare with a direct calculation on a calculator.

$R_1 = 100$ $\quad R_2 = 200$ $\quad R_3 = 300$

$\Delta R_1 = 1$ $\quad \Delta R_2 = 2$ $\quad \Delta R_3 = 4$.

→ 求 $\frac{\partial R}{\partial R_1}$

$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}$

$\frac{\partial}{\partial R_1}\left(\frac{1}{R}\right) = \frac{\partial}{\partial R_1}\left(\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}\right) = -\frac{1}{(R_1)^2}$

$-\frac{1}{R^2}\frac{\partial R}{\partial R_1} = -\frac{1}{R_1^2}$ $\qquad \frac{\partial R}{\partial R_1} = \frac{R^2}{R_1^2}$ $\quad$ ($\frac{\partial R}{\partial R_2}$, $\frac{\partial R}{\partial R_3}$ 同理)

$\Delta R \approx \frac{\partial R}{\partial R_1}(R_1, R_2, R_3)\Delta R_1 + \frac{\partial R}{\partial R_2}(R_1, R_2, R_3) + \frac{\partial R}{\partial R_3}(R_1, R_2, R_3)$

更清晰的: $R = \dfrac{1}{\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}}$

19. Find all planes which are tangent to the surface $z = 1 - x^2 - y^2$, and contain the line passing through the points $(1, 0, 2)$ and $(0, 2, 2)$.

→ find partial derivatives 代入 (a, b)

$f_x = -2x$ $\qquad f_x(a, b) = -2a$

$f_y = -2y$ $\qquad f_y(a, b) = -2b$

→ general equation of tangent plane

$z = 1 - a^2 - b^2 - 2a(x - a) - 2b(x - b)$

→ 用 b 代替 a

代入 $(1, 0, 2)$ $(0, 2, 2)$

$\begin{cases} 2 = 1 + a^2 + b^2 - 2a \\ 2 = 1 + a^2 + b^2 - 4b \end{cases} \Rightarrow a = 2b$

$z = 5b^2 - 4bx - 2by + 1$ $\quad b \in \mathbb{R}$

Chapter 5

Differentiable Functions

5.1 Definition of Differentiability

Now, our goal is to extend the concept of differentiability for functions of one variable to functions of two variables. For a function of one variable, we saw that a function $g(x)$ is differentiable at $x = a$ if $g'(a)$ exists. From this, it is natural to wonder if the existence of partial derivatives is enough to define the concept of differentiability for $f(x, y)$. Unfortunately, it isn't. We saw in Example 4.1.3 that the concept of partial derivatives does not match with the concept of differentiability from Calculus 1. In particular, we saw a function whose partial derivatives exist at $(0, 0)$ even though the function is not continuous at $(0, 0)$.

To define the concept of differentiability for $f(x, y)$, we want to ensure that it has the same properties as the concept of differentiability for functions of one variable. In Calculus 1, we saw that if $g(x)$ is differentiable at $x = a$, then graph of $g(x)$ is 'smooth' at $x = a$ (no cusps or jumps) and that the linear approximation is a good approximation. In particular, if we define the error in the linear approximation to be

$$R_{1,a}(x) = g(x) - L_a(x)$$

then we get the following theorem.

THEOREM 1 If $g'(a)$ exists, then $\lim_{x \to a} \frac{|R_{1,a}(x)|}{|x - a|} = 0$ where

$$R_{1,a}(x) = g(x) - L_a(x) = g(x) - g(a) - g'(a)(x - a)$$

Proof: We have

$$\frac{|R_{1,a}(x)|}{|x - a|} = \left|\frac{g(x) - g(a) - g'(a)(x - a)}{x - a}\right| = \left|\frac{g(x) - g(a)}{x - a} - g'(a)\right|$$

The result follows from taking the limit as $x \to a$ (details left as an exercise). □

Theorem 1 says that the error $R_{1,a}(x)$ tends to zero faster than the displacement $|x-a|$. Moreover, it can be shown that if one replaces the tangent line $y = L_a(x)$ by any other straight line $y = g(a) + m(x - a)$ through the point $(a, g(a))$, the error will not satisfy the conclusion of the theorem. Thus, the property

$$\lim_{x\to a} \frac{|R_{1,a}(x)|}{|x-a|} = 0$$

characterizes the tangent line at $(a, g(a))$ as the best straight line approximation to the graph $y = g(x)$ near $(a, g(a))$.

Therefore, to define differentiability for a function of two variables, we match this definition.

DEFINITION

Differentiable

A function $f(x, y)$ is **differentiable** at (a, b) if

$$\lim_{(x,y)\to(a,b)} \frac{|R_{1,(a,b)}(x, y)|}{\|(x, y) - (a, b)\|} = 0$$

where

$$R_{1,(a,b)}(x, y) = f(x, y) - L_{(a,b)}(x, y)$$

As in the one dimensional case, we can prove that the only tangent plane $z = f(a, b)+c(x-a)+d(y-b)$ through the point $(a, b, f(a, b)$ that has this property is $z = L_{(a,b)}(x, y)$.

THEOREM 2

If a function $f(x, y)$ satisfies

$$\lim_{(x,y)\to(a,b)} \frac{|f(x, y) - f(a, b) - c(x - a) - d(y - b)|}{\|(x, y) - (a, b)\|} = 0$$

then $c = f_x(a, b)$ and $d = f_y(a, b)$.

Proof: Since

$$\lim_{(x,y)\to(a,b)} \frac{|f(x, y) - f(a, b) - c(x - a) - d(y - b)|}{\|(x, y) - (a, b)\|} = 0$$

the limit is 0 along any path. Therefore, along the path along $y = b$, we get

$$\begin{aligned}
0 &= \lim_{x\to a} \frac{|f(x, b) - f(a, b) - c(x - a) - d(b - b)|}{\|(x, b) - (a, b)\|} \\
&= \lim_{x\to a} \frac{|f(x, b) - f(a, b) - c(x - a)|}{|x - a|} \\
&= \lim_{x\to a} \left| \frac{f(x, b) - f(a, b)}{x - a} - c \right| \\
&= f_x(a, b) - c \\
c &= f_x(a, b)
\end{aligned}$$

Similarly, approaching along $x = a$ we get that $d = f_y(a, b)$. □

This implies that the tangent plane gives the best linear approximation to the graph $z = f(x, y)$ near (a, b). Moreover, it tells us that the linear approximation is a "good approximation" if and only if f is differentiable at (a, b).

REMARK

Observe that for the linear approximation to exist at (a, b) both partial derivatives of f must exist at (a, b). However, both partial derivatives existing does not guarantee that f will be differentiable. We say that the partial derivatives of f existing at (a, b) is necessary, but not sufficient.

EXAMPLE 1 Determine whether $f(x, y) = \sqrt{|xy|}$ is differentiable at $(0, 0)$.

Solution: We first need to find $L_{(0,0)}(x, y)$. Hence, we need to find the partial derivatives at $(0, 0)$. We have

$$\begin{aligned} f_x(0,0) &= \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} = \lim_{h\to 0} \frac{0-0}{h} = 0 \\ f_y(0,0) &= \lim_{h\to 0} \frac{f(0,h) - f(0,0)}{h} = \lim_{h\to 0} \frac{0-0}{h} = 0 \end{aligned}$$

Thus, both partial derivatives exist at $(0, 0)$ and, since $f(0, 0) = 0$, the linear approximation is

$$L_{(0,0)}(x, y) = 0$$

The error in the linear approximation is

$$R_{1,(0,0)}(x, y) = f(x, y) - L_{(0,0)}(x, y) = \sqrt{|xy|}$$

and the magnitude of the displacement is $\|(x, y) - (0, 0)\| = \sqrt{x^2 + y^2}$.

Therefore,

$$\frac{|R_{1,(0,0)}(x, y)|}{\|(x, y) - (0, 0)\|} = \frac{\sqrt{|xy|}}{\sqrt{x^2 + y^2}}, \qquad \text{for } (x, y) \neq (0, 0)$$

We must determine whether

$$\lim_{(x,y)\to(0,0)} \frac{\sqrt{|xy|}}{\sqrt{x^2 + y^2}} = 0 \tag{5.1}$$

As we saw with continuity, to prove the limit does not equal to 0, we do not need to prove that the limit does not exist. We just need to prove that there exists a single path that gives a limit other than 0. In this case, approaching along the line $y = x$ gives

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

so that

$$\lim_{(x,y)\to(0,0)} \frac{|R_{1,(0,0)}(x, y)|}{\|(x, y) - (0, 0)\|} \neq 0$$

It follows that (5.1) is false. Thus, by definition, the given function f is not differentiable at $(0, 0)$.

Observe that in this example we have that the partial derivatives at $(0,0)$ both exist, but

$$\lim_{(x,y)\to(0,0)} \frac{|R_{1,(0,0)}(x,y)|}{\|(x,y)-(0,0)\|} \neq 0$$

So, the plane $z = L_{(0,0)}(x,y) = 0$ does not give a good approximation to the surface $z = \sqrt{|xy|}$ near the origin. This can be explained geometrically. The vertical plane $y = x$ intersects the surface $z = \sqrt{|xy|}$ in the curve $z = |x|$ which has a corner at $x = 0$ and hence no tangent line. This means that the surface is not "smooth" at $(0,0,0)$, and hence the plane $z = L_{(0,0)}(x,y) = 0$ cannot be interpreted as a tangent plane.

EXAMPLE 2 Show that $f(x,y) = x^2 + y^2$ is differentiable at $(a,b) = (1,0)$.

Solution: We have $f_x = 2x$ and $f_y = 2y$, so at $(1,0)$ we get $f_x(1,0) = 2$ and $f_y(1,0) = 0$. Thus, we have

$$\begin{aligned} L_{(1,0)}(x,y) &= f(1,0) + f_x(1,0)(x-1) + f_y(1,0)(y-0) \\ &= 1 + 2(x-1) + 0(y-0) = 1 + 2(x-1) \end{aligned}$$

The error in the linear approximation is

$$\begin{aligned} R_{1,(1,0)}(x,y) &= f(x,y) - L_{(1,0)}(x,y) \\ &= x^2 + y^2 - (1 + 2(x-1)) \\ &= x^2 - 2x + 1 + y^2 = (x-1)^2 + y^2 \end{aligned}$$

Therefore,

$$\frac{|R_{1,(1,0)}(x,y)|}{\|(x,y)-(1,0)\|} = \frac{(x-1)^2 + y^2}{\sqrt{(x-1)^2 + y^2}} = \sqrt{(x-1)^2 + y^2}$$

Hence,

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

by the Continuity Theorems. So, $f(x,y)$ is differentiable at $(1,0)$.

EXAMPLE 3 Determine whether $f(x,y) = \begin{cases} \dfrac{x^2 y}{x^2 + y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$ is differentiable at $(0,0)$.

Solution: We have

$$\begin{aligned} f_x(0,0) &= \lim_{h\to 0} \frac{f(0+h,0) - f(0,0)}{h} = \lim_{h\to 0} \frac{\frac{h^2(0)}{h^2+0^2} - 0}{h} = 0 \\ f_y(0,0) &= \lim_{h\to 0} \frac{f(0,0+h) - f(0,0)}{h} = \lim_{h\to 0} \frac{\frac{0^2(h)}{0^2+h^2} - 0}{h} = 0 \end{aligned}$$

So, the error in the linear approximation is

$$R_{1,(0,0)}(x,y) = f(x,y) - f(0,0) - f_x(0,0)(x-0) - f_y(0,0)(y-0) = \frac{x^2 y}{x^2 + y^2}$$

For f to be differentiable at $(0,0)$ we need $\lim_{(x,y)\to(0,0)} \frac{|R_{1,(0,0)}(x,y)|}{\sqrt{x^2+y^2}} = 0$. But, observe that if we approach the limit along $y = x$, we get

$$\begin{aligned}\lim_{x\to 0} \frac{|R_{1,(0,0)}(x,x)|}{\sqrt{x^2+x^2}} &= \lim_{x\to 0} \frac{|x^3|}{(x^2+x^2)^{3/2}} \\ &= \lim_{x\to 0} \frac{|x^3|}{(2x^2)^{3/2}} \\ &= \lim_{x\to 0} \frac{|x^3|}{2^{3/2}|x^3|} \\ &= \lim_{x\to 0} \frac{1}{2^{3/2}} = \frac{1}{2^{3/2}}\end{aligned}$$

Therefore, the limit can not equal 0 and hence f is not differentiable at $(0,0)$.

EXAMPLE 4 Determine whether $g(x,y) = \begin{cases} \dfrac{x^2y^2}{x^2+y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$ is differentiable at $(0,0)$.

Solution: We have

$$\begin{aligned} g_x(0,0) &= \lim_{h\to 0} \frac{g(0+h,0) - g(0,0)}{h} = \lim_{h\to 0} \frac{\frac{h^2(0)^2}{h^2+0^2} - 0}{h} = 0 \\ g_y(0,0) &= \lim_{h\to 0} \frac{g(0,0+h) - g(0,0)}{h} = \lim_{h\to 0} \frac{\frac{0^2(h)^2}{0^2+h^2} - 0}{h} = 0 \end{aligned}$$

So, the error in the linear approximation is

$$R_{1,(0,0)}(x,y) = g(x,y) - g(0,0) - g_x(0,0)(x-0) - g_y(0,0)(y-0) = \frac{x^2y^2}{x^2+y^2}$$

For g to be differentiable at $(0,0)$ we need $\lim_{(x,y)\to(0,0)} \frac{|R_{1,(0,0)}(x,y)|}{\sqrt{x^2+y^2}} = 0$. If we approach the limit along $y = mx$, we get

$$\begin{aligned}\lim_{x\to 0} \frac{|R_{1,(0,0)}(x,mx)|}{\sqrt{x^2+(mx)^2}} &= \lim_{x\to 0} \frac{m^2x^4}{(x^2+m^2x^2)^{3/2}} \\ &= \lim_{x\to 0} \frac{m^2x^4}{([1+m^2]x^2)^{3/2}} \\ &= \lim_{x\to 0} \frac{m^2x^4}{(1+m^2)^{3/2}|x^3|} \\ &= \lim_{x\to 0} \frac{m^2|x|}{(1+m^2)^{3/2}} = 0\end{aligned}$$

So, perhaps the limit exists. We try to apply the Squeeze Theorem. We consider

$$\begin{aligned}\left|\frac{x^2y^2}{(x^2+y^2)^{3/2}} - 0\right| &\leq \frac{(x^2+y^2)(x^2+y^2)}{(x^2+y^2)^{3/2}} \\ &= \frac{(x^2+y^2)^2}{(x^2+y^2)^{3/2}} \\ &= (x^2+y^2)^{1/2}\end{aligned}$$

Since

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

by the Continuity Theorems, we get by the Squeeze Theorem that

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

Hence, g is differentiable at $(0,0)$.

REMARK

In Example 3.1.1 we showed that $f(x,y) = \begin{cases} \dfrac{x^2y}{x^2+y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$ is continuous at $(0,0)$. So, this is an example of a function that is continuous but not differentiable at a point. In the next section, we will prove that if a function is differentiable at a point, then it must be continuous at that point to match what we saw in Calculus 1.

EXERCISE 1 Prove that

$$f(x,y) = \begin{cases} \dfrac{x^3}{x^2+y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$$

is not differentiable at $(0,0)$.

EXERCISE 2 Prove that $f(x,y) = |xy|$ is differentiable at $(0,0)$.

EXERCISE 3 Prove that $f(x,y) = |xy|$ is not differentiable at $(0,1)$.

We can now give a formal definition of the tangent plane of $z = f(x,y)$.

DEFINITION
Tangent Plane

Consider a function $f(x, y)$ which is differentiable at (a, b). The **tangent plane** of the surface $z = f(x, y)$ at $(a, b, f(a, b))$ is the graph of the linearization. That is, the tangent plane is given by

$$z = f(a, b) + \frac{\partial f}{\partial x}(a, b)(x - a) + \frac{\partial f}{\partial y}(a, b)(y - b)$$

Since f is assumed to be differentiable at (a, b), by Theorem 2, the tangent plane is the plane that best approximates the surface near the point $(a, b, f(a, b))$. In this case, we say that at the point $(a, b, f(a, b))$ the surface $z = f(x, y)$ is **smooth**.

EXERCISE 4

Invent a function $f(x, y)$ whose graph $z = f(x, y)$ is not smooth at $(1, 2, f(1, 2))$. That is, invent a function which is not differentiable at $(1, 2)$.

5.2 Differentiability and Continuity

Recall from Calculus 1 that if $g(x)$ is differentiable at $x = a$, then g is continuous at a. We now prove that this result also holds for scalar functions $f(x, y)$.

THEOREM 1

If $f(x, y)$ is differentiable at (a, b), then f is continuous at (a, b).

Proof: The error $R_{1,(a,b)}(x, y)$ is defined by

$$R_{1,(a,b)}(x, y) = f(x, y) - L_{(a,b)}(x, y)$$

Using the definition of $L_{(a,b)}(x, y)$, this equation can be rearranged to read

$$f(x, y) = f(a, b) + \nabla f(a, b) \cdot (x - a, y - b) + R_{1,(a,b)}(x, y) \tag{5.2}$$

We can write

$$R_{1,(a,b)}(x, y) = \frac{R_{1,(a,b)}(x, y)}{\|(x, y) - (a, b)\|}\|(x, y) - (a, b)\|, \quad \text{for} \quad (x, y) \neq (a, b)$$

Since f is differentiable and by the Limit Theorems, we get

$$\lim_{(x,y)\to(a,b)} R_{1,(a,b)}(x, y) = 0$$

It now follows from equation (5.2) that

$$\lim_{(x,y)\to(a,b)} f(x, y) = f(a, b) + 0 + 0 = f(a, b)$$

and so by definition, f is continuous at (a, b). $\square$

EXERCISE 1 Suppose that $f(x, y)$ is not continuous at (a, b). Can you draw a conclusion about whether f is differentiable at (a, b)?

EXERCISE 2 Give an example of a function $f(x, y)$ that is continuous but not differentiable at (a, b). This shows that the converse of Theorem 1 is not true.

5.3 Continuous Partial Derivatives and Differentiability

We need an efficient way of proving that a given function f is differentiable at a typical point. In this section, we present a theorem for this purpose, which states that if the partial derivatives of $f(x, y)$ are continuous at (a, b), then f is differentiable at (a, b).

To prove this theorem, we will require an extremely important theorem from Calculus 1, the Mean Value Theorem.

THEOREM 1 **(Mean Value Theorem)**

If $f(t)$ is continuous on the closed interval $[t_1, t_2]$ and f is differentiable on the open interval (t_1, t_2), then there exists $t_0 \in (t_1, t_2)$ such that

$$f(t_2) - f(t_1) = f'(t_0)(t_2 - t_1)$$

THEOREM 2 If $\dfrac{\partial f}{\partial x}$ and $\dfrac{\partial f}{\partial y}$ are continuous at (a, b), then $f(x, y)$ is differentiable at (a, b).

Proof: We derive an expression for the error $R_{1,(a,b)}(x, y)$, given by

$$R_{1,(a,b)}(x, y) = f(x, y) - f(a, b) - f_x(a, b)(x - a) - f_y(a, b)(y - b) \tag{5.3}$$

Since f_x and f_y are continuous then f_x and f_y exist in some neighborhood $B(a, b)$. For $(x, y) \in B(a, b)$, we write

$$f(x, y) - f(a, b) = \Big[f(x, y) - f(a, y)\Big] + \Big[f(a, y) - f(a, b)\Big] \tag{5.4}$$

by adding and subtracting $f(a, y)$. The Mean Value Theorem can be applied to each bracket, since one variable is held fixed, and the partial derivatives are assumed to exist. For the first bracket:

$$f(x, y) - f(a, y) = f_x(\overline{x}, y)(x - a)$$

where $\overline{x}$ lies between a and x. By adding and subtracting $f_x(a, b)(x - a)$, we obtain

$$f(x, y) - f(a, y) = f_x(a, b)(x - a) + A(x - a) \tag{5.5}$$

where

$$A = f_x(\overline{x}, y) - f_x(a, b) \tag{5.6}$$

Similarly for the second bracket:

$$\begin{aligned} f(a, y) - f(a, b) &= f_y(a, \overline{y})(y - b) \\ &= f_y(a, b)(y - b) + B(y - b) \end{aligned} \tag{5.7}$$

where

$$B = f_y(a, \overline{y}) - f_y(a, b) \tag{5.8}$$

and $\overline{y}$ lies between b and y.

Substitute equations (5.5) and (5.7) into (5.4) and then substitute equation (5.4) into (5.3) to obtain

$$R_{1,(a,b)}(x, y) = A(x - a) + B(y - b)$$

where A and B are given by equations (5.6) and (5.8). It follows by the triangle inequality that

$$\begin{aligned} \frac{|R_{1,(a,b)}(x, y)|}{\|(x, y) - (a, b)\|} &\leq \frac{|A||x - a|}{\sqrt{(x - a)^2 + (y - b)^2}} + \frac{|B||y - b|}{\sqrt{(x - a)^2 + (y - b)^2}} \\ &\leq |A| + |B| \end{aligned} \tag{5.9}$$

We can now apply the Squeeze Theorem with $L = 0$ and $B(x, y) = |A| + |B|$.

As $(x, y) \to (a, b)$, it follows that

$$(\overline{x}, y) \to (a, b) \quad \text{and} \quad (a, \overline{y}) \to (a, b)$$

Since f_x and f_y are continuous at (a, b), it follows from equations (5.6) and (5.8) that

$$\lim_{(x,y)\to(a,b)} A = 0 \quad \text{and} \quad \lim_{(x,y)\to(a,b)} B = 0$$

Equation (5.9) and the Squeeze Theorem now imply

$$\lim_{(x,y)\to(a,b)} \frac{|R_{1,(a,b)}(x, y)|}{\|(x, y) - (a, b)\|} = 0$$

so that f is differentiable at (a, b), by definition. □

REMARK

The converse of Theorem 2 is *not* true. That is, $f(x, y)$ being differentiable at (a, b) does not imply that f_x and f_y are both continuous at (a, b).

EXAMPLE 1

Determine at which points $f(x, y) = (x^2 + y^2)^{2/3}$ is differentiable.

Solution: By differentiation

$$\frac{\partial f}{\partial x} = \frac{4x}{3(x^2 + y^2)^{1/3}}, \quad \text{for} \quad (x, y) \neq (0, 0)$$

By inspection, using the Continuity Theorems, $\dfrac{\partial f}{\partial x}$ is continuous for all $(x, y) \neq (0, 0)$. By symmetry, the same conclusion holds for $\dfrac{\partial f}{\partial y}$. It follows from Theorem 2 that f is differentiable for all $(x, y) \neq (0, 0)$.

At the point $(0, 0)$, it is not clear whether the partial derivatives exist and one has to use the definition of partial derivative. Then one has to use the definition of differentiable function, as in Example 5.1.1. The conclusion is that f is differentiable at $(0, 0)$.

EXERCISE 1

Prove that $f(x, y) = (x^2 + y^2)^{2/3}$ is differentiable at $(0, 0)$.

EXERCISE 2

Prove that if $f(x, y) \in C^2$ at (a, b), then f is continuous at (a, b).

Summary

Theorem 2 makes it easy to prove that a function f is differentiable at a typical point. One simply differentiates f to obtain the partial derivatives f_x, f_y, and then checks that the partials are continuous functions by inspection, referring to the Continuity Theorems, as in Section 3.2. It is only necessary to use the definition of a differentiable function at an exceptional point.

Generalization

The definition of a differentiable function and theorems 1 and 2 are valid for functions of n variables. The only change is that there are n partial derivatives,

$$\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}$$

5.4 Linear Approximation Revisited

The error in the linear approximation for $f(x, y)$ is defined by

$$R_{1,(a,b)}(x, y) = f(x, y) - L_{(a,b)}(x, y)$$

where

$$L_{(a,b)}(x, y) = f(a, b) + \nabla f(a, b) \cdot ((x, y) - (a, b))$$

It is convenient to rearrange the definition of $R_{1,(a,b)}(x, y)$ to read

$$f(x, y) = f(a, b) + \nabla f(a, b) \cdot (x - a, y - b) + R_{1,(a,b)}(x, y) \tag{5.10}$$

The linear approximation

$$f(x, y) \approx f(a, b) + \nabla f(a, b) \cdot (x - a, y - b) \tag{5.11}$$

for (x, y) sufficiently close to (a, b), arises if one neglects the error term. In general, one has no information about $R_{1,(a,b)}(x, y)$, and so it is not clear whether the approximation is reasonable. However, Theorem 2 provides an important piece of information about $R_{1,(a,b)}(x, y)$, namely that if the partial derivatives of f are continuous at (a, b), then f is differentiable and hence

$$\lim_{(x,y)\to(a,b)} \frac{|R_{1,(a,b)}(x, y)|}{\|(x, y) - (a, b)\|} = 0$$

In this case, the approximation (5.11) is reasonable for (x, y) sufficiently close to (a, b), and we say that $L_{(a,b)}(x, y)$ is a good approximation of $f(x, y)$ near (a, b).

EXAMPLE 1 Discuss the validity of the approximation

$$(xy)^{1/3} \approx 2 + \frac{1}{3}(x - 2) + \frac{1}{6}(y - 4)$$

Solution: Let $f(x, y) = (xy)^{1/3}$. By differentiation,

$$\nabla f(x, y) = \left(\frac{1}{3}x^{-\frac{2}{3}}y^{\frac{1}{3}}, \quad \frac{1}{3}x^{\frac{1}{3}}y^{-\frac{2}{3}}\right)$$

so $\nabla f(2, 4) = \left(\frac{1}{3}, \frac{1}{6}\right)$. With $(a, b) = (2, 4)$, equation (5.10) becomes

$$(xy)^{\frac{1}{3}} = 2 + \frac{1}{3}(x - 2) + \frac{1}{6}(y - 4) + R_{1,(2,4)}(x, y)$$

Using the Continuity Theorems we see that f has continuous partials at the point $(2, 4)$. Thus,

$$\lim_{(x,y)\to(2,4)} \frac{|R_{1,(2,4)}(x, y)|}{\sqrt{(x - 2)^2 + (y - 4)^2}} = 0$$

It follows that for (x, y) sufficiently close to $(2, 4)$, we may neglect $R_{1,(2,4)}(x, y)$. Thus,

$$(xy)^{\frac{1}{3}} \approx 2 + \frac{1}{3}(x - 2) + \frac{1}{6}(y - 4)$$

gives a good approximation for (x, y) sufficiently close to $(2, 4)$.

EXERCISE 1 Discuss the validity of the approximation

$$\sqrt{1 + 3\tan x + \sin y} \approx 2 + \frac{3}{2}\left(x - \frac{\pi}{4}\right) + \frac{1}{4}y$$

Note that approximation is a recurring theme in calculus, and the equation

$$f(\mathbf{x}) = f(\mathbf{a}) + \nabla f(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a}) + R_{1,\mathbf{a}}(\mathbf{x})$$

is of fundamental importance. In Chapter 8, we shall find out more about the error term $R_{1,\mathbf{a}}(\mathbf{x})$ in terms of the second partial derivatives.

EXERCISE 2 For each of the following either give an example to prove that the statement is false, or justify why the statement is true.

(a) If f is not continuous at $(0, 0)$, then f_x and f_y cannot both be continuous at $(0, 0)$.

(b) If f is continuous at $(0, 0)$ and both $f_x(0, 0)$ and $f_y(0, 0)$ exist, then f is differentiable at $(0, 0)$.

(c) If f is not differentiable at $(0, 0)$, then at least one of $f_x(0, 0)$ or $f_y(0, 0)$ does not exist.

(d) If f is not differentiable at $(0, 0)$, then f is not continuous at $(0, 0)$.

(e) If f is differentiable at $(0, 0)$, then both $f_x(0, 0)$ and $f_y(0, 0)$ exist.

(f) If f is differentiable at $(0, 0)$, then f_x and f_y are both continuous at $(0, 0)$.

Chapter 5 Problem Set

1. Prove that $f(x, y) = x(|y| - 1)$ is differentiable at $(0, 0)$.

2. Prove that $f(x, y) = x|y-1|$ is not differentiable at $(1, 1)$.

3. (a) Sketch the surface $z = |x-y|$ in $\mathbb{R}^3$. At what points does the surface not have a tangent plane?
   (b) Verify that the partial derivatives of the function $f(x, y) = |x - y|$ do not exist at the points found in (a).

   Note: This implies that f is not differentiable at these points, as expected from (a).

4. For each function
   (a) Use the definition to determine whether f is differentiable at $(0, 0)$.
   (b) On the basis of your answer in (a), can you use one of the theorems to draw a conclusion concerning the continuity of f at $(0, 0)$?
   (c) On the basis of your answer in (a), can you use one of the theorems to draw a conclusion concerning the continuity of f_x and f_y at $(0, 0)$?

   (i) $f(x, y) = (xy)^{2/3}$
   (ii) $f(x, y) = (xy)^{1/3}$
   (iii) $f(x, y) = |x|^{\frac{1}{2}}|y|^{\frac{3}{2}}$
   (iv) $f(x, y) = \begin{cases} \dfrac{x^3 + y^4}{x^2 + y^2} & \text{for } (x, y) \neq (0, 0) \\ 0 & \text{for } (x, y) = (0, 0) \end{cases}$

5. For each of the following functions, determine if f is differentiable at $(0, 0)$.
   (a) $f(x, y) = \begin{cases} \dfrac{x^4 + y^4}{x^2 + y^2} + 1 & \text{if } (x, y) \neq (0, 0) \\ 1 & \text{if } (x, y) = (0, 0) \end{cases}$
   (b) $f(x, y) = \begin{cases} \dfrac{x|y|}{\sqrt{x^2 + y^2}} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{if } (x, y) = (0, 0) \end{cases}$

6. Let $f(x, y) = \begin{cases} \dfrac{x^3 - y^4}{x^2 + y^2} + 1 & \text{if } (x, y) \neq (0, 0) \\ 1 & \text{if } (x, y) = (0, 0) \end{cases}$.
   Determine all points where f is differentiable.

7. Let $f(x, y) = \begin{cases} \dfrac{xy^2 + y^3}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{if } (x, y) = (0, 0) \end{cases}$
   (a) Prove that f is continuous at $(0, 0)$.
   (b) Determine all points where f is differentiable.

8. Determine whether the functions in #4 are differentiable at $(0, a)$, $a \neq 0$.
   Hint: Does $f_x(0, a)$ exist? Consider the cross-section $y = a$ to get a geometric interpretation.

9. Determine all points where f is differentiable.
   (a) $f(x, y) = \begin{cases} \dfrac{x^3 + y^3}{x^2 + y^2} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{if } (x, y) = (0, 0) \end{cases}$
   (b) $g(x, y) = |x|^{1/2}$

10. (a) Invent a function $f(x, y)$ which is continuous on $\mathbb{R}^2$ but not differentiable at $(1, 2)$. Sketch the surface $z = f(x, y)$.
    (b) Invent a function $f(x, y)$ which is continuous on $\mathbb{R}^2$ but not differentiable at all points of the circle $x^2 + y^2 = 1$. Sketch the surface $z = f(x, y)$.

11. (a) Find a function $f(x, y)$ such that $f(x, y)$ is continuous at $(0, 0)$, but $f_x(0, 0)$ and $f_y(0, 0)$ both do not exist. Justify your answer.
    (b) Find a function $g(x, y)$ such that $g_x(0, 0)$ and $g_y(0, 0)$ both exist at $(0, 0)$, but $g(x, y)$ is not continuous at $(0, 0)$. Justify your answer.

12. Let $f(x, y) = \begin{cases} \dfrac{x^4 y^2}{x^2 + y^2} + 1 & \text{if } (x, y) \neq (0, 0) \\ 1 & \text{if } (x, y) = (0, 0). \end{cases}$
    (a) Determine $f_x(0, 0)$ and $f_y(0, 0)$.
    (b) Determine $f_x(x, y)$ for all $(x, y) \neq (0, 0)$.
    (c) Determine if f_x is continuous at $(0, 0)$.

13. Consider the theorem: If $f(x, y)$ is differentiable at (a, b) then f is continuous at (a, b). Give a counterexample to show that the converse of the theorem is false.

14. Prove that if f_{xx}, f_{xy}, f_{yx} and f_{yy} are continuous at (a, b), then f_x, f_y and f are continuous at (a, b).
    Hint: Apply the theorems relating to differentiability.

15. * Let $f(x, y) = |x|^r |y|^s$, where r, s are positive numbers.
    (a) For what values of r and s is f differentiable at (0,0)?
    (b) For what values of r and s is f differentiable on $\mathbb{R}^2$?

16. * Prove that if f satisfies $|f(x, y)| \leq x^2 + y^2$ for all $(x, y) \in \mathbb{R}^2$, then f is differentiable at (0,0).

$f_x(0,y) = -y \qquad f_y(x,0) = x$

$f_{xy}(0,0) = -1 \neq f_{yx}(0,0) = 1$

1. Prove that $f(x, y) = x(|y| - 1)$ is differentiable at $(0, 0)$.

$\to f_x = |y| - 1 \qquad f_x(0,0) = -1$

$f_y = x\frac{y}{|y|} \qquad f_y(0,0) = \lim_{h\to 0}\frac{f(0,h)-f(0,0)}{h}$

$= \lim_{h\to 0}\frac{0-0}{h}$

$= 0$

$\to L_{(0,0)}(x,y) = f(0,0) - 1(x-0) + 0(y-0)$

$= 0 - x$

$= -x$

$\to R_{1,(0,0)}(x,y) = f(x,y) - L_{(0,0)}(x,y)$

$= x(|y|-1) - (-x)$

$= x|y|$

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

approach along $y = mx$

$\lim_{(x,y)\to(0,0)}\frac{|x||y|}{\sqrt{x^2+y^2}} = \lim_{x\to 0}\frac{|x||mx|}{\sqrt{x^2+m^2x^2}}$

$= \lim_{x\to 0}\frac{|x||mx|}{|x|\sqrt{1+m^2}}$

$= \lim_{x\to 0}\frac{|m||x|}{\sqrt{1+m^2}}$

$= 0$

Let $f(x,y) = \frac{|x||y|}{\sqrt{x^2+y^2}}$, $L = 0$

$|f(x,y) - L| = \left|\frac{|x||y|}{\sqrt{x^2+y^2}} - 0\right|$

$\le \frac{2|x||y|}{\sqrt{x^2+y^2}}$

$\le \frac{x^2+y^2}{\sqrt{x^2+y^2}}$

$= \sqrt{x^2+y^2}$

$0 \le \lim_{(x,y)\to(0,0)}\frac{|R_{1(a,b)}(x,y)|}{\|(x,y)-(a,b)\|} \le \lim_{(x,y)\to(0,0)}\sqrt{x^2+y^2} = 0$

$\therefore \lim_{(0,0)} \to 0$

$\therefore$ is diff

2. Prove that $f(x, y) = x|y-1|$ is not differentiable at $(1, 1)$.

$\to f_x = |y-1| \qquad f_x(1,1) = 0$

$f_y = x\frac{y-1}{|y-1|} \qquad f_y(1,1) = \lim_{h\to 0}\frac{f(0,h)-f(0,0)}{h}$

$= \lim_{h\to 0}\frac{0}{0} \quad \therefore$ DNE

$\therefore$ not diff

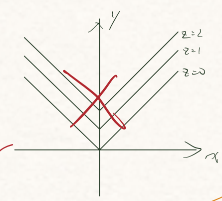
3. (a) Sketch the surface $z = |x-y|$ in $\mathbb{R}^3$. At what points does the surface not have a tangent plane?

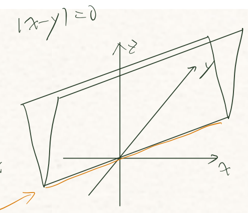
(b) Verify that the partial derivatives of the function $f(x, y) = |x-y|$ do not exist at the points found in (a).

Note: This implies that f is not differentiable at these points, as expected from (a).

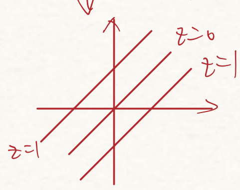
$f_x(x,y)$

(a)





$(k, k, 0) \quad (x = y$ 时$)$



(b) $f_x(k,k) = \lim_{h\to 0}\frac{f(k+h,k)-f(k,k)}{h}$

$= \lim_{h\to 0}\frac{|k+h-k|-|k-k|}{h}$

$= \lim_{h\to 0}\frac{|h|}{h}$

$= \pm 1 \qquad$ DNE

?

(c) If $f(x, y)$ is C^2 at (a, b) then $f(x, y)$ is differentiable at (a, b).

TRUE FALSE

second derivative is cont. function

4. For each function

(a) Use the definition to determine whether f is differentiable at $(0,0)$.

(b) On the basis of your answer in (a), can you use one of the theorems to draw a conclusion concerning the continuity of f at $(0,0)$?

(c) On the basis of your answer in (a), can you use one of the theorems to draw a conclusion concerning the continuity of f_x and f_y at $(0,0)$?

(i) $f(x,y) = (xy)^{2/3}$

(ii) $f(x,y) = (xy)^{1/3}$

(iii) $f(x,y) = |x|^{\frac{1}{2}}|y|^{\frac{3}{2}}$

(iv) $f(x,y) = \begin{cases} \dfrac{x^3+y^4}{x^2+y^2} & \text{for } (x,y) \neq (0,0) \\ 0 & \text{for } (x,y) = (0,0) \end{cases}$

(i) (a) $f_x(0,0) = 0$

$f_y(0,0) = 0$

$L_{(0,0)}(x,y) = f(0,0) = 0$

$R_{1,(0,0)}(x,y) = f(x,y) - L_{(0,0)}(x,y)$
$= (xy)^{\frac{2}{3}}$

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

(b) cont.
Since diff → cont

(c) 未知

(ii) (a) $f_x(x,y) = \frac{1}{3}x^{-\frac{2}{3}} \cdot y^{\frac{1}{3}}$

approach along $y = mx$

$\lim_{(0,0)} f_x(x,y) = \lim_{x\to 0} \frac{1}{3}x^{-\frac{2}{3}}(mx)^{\frac{1}{3}} = \frac{1}{3}m^{\frac{1}{3}} \lim_{x\to 0} x^{-\frac{1}{3}} = \infty$

$f_x(0,0)$ & $f_y(0,0)$ DNE

(b) don't know whether cont

(c) 未知

(iii) (a) diff

(b) cont diff → cont

(c) 未知

(iv) (a) $f_x = \lim_{h\to 0} \dfrac{\frac{h^3}{h^2}}{h} = 1$

$f_y = 1$

$L = 0 + x + y$

→ approach along $y = mx$.

lim depend on m

∴ $f(x,y)$ is diff except at $(0,0)$

(b) N

(c) N

5. For each of the following functions, determine if f is differentiable at $(0,0)$.

(a) $f(x,y) = \begin{cases} \dfrac{x^4+y^4}{x^2+y^2} + 1 & \text{if } (x,y) \neq (0,0) \\ 1 & \text{if } (x,y) = (0,0) \end{cases}$

(b) $f(x,y) = \begin{cases} \dfrac{x|y|}{\sqrt{x^2+y^2}} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$

(a) → $f_x = \dfrac{(4x^3+y^4)(x^2+y^2) - (x^4+y^4)(2x+y^2)}{(x^2+y^2)^2}$ $\forall (x,y) \neq (0,0)$

$f_y = \dfrac{(x^4+4y^3)(x^2+y^2) - (x^4+y^4)(x^2+2y)}{(x^2+y^2)^2}$ $\forall (x,y) \neq (0,0)$

by continuity theorems, these function cont $\forall (x,y) \neq (0,0)$

→ $f_x(0,0) = \lim_{h\to 0} \dfrac{h^4 - 0^4}{h^2 - 0} = 0$

$f_y(0,0) = \lim_{h\to 0} \dfrac{h^4}{h^2} = 0$

$L_{(0,0)}(x,y) = 1$

$R_{1(0,0)}(x,y) = \dfrac{x^4+y^4}{x^2+y^2} + 1 - 1 = \dfrac{x^4+y^4}{x^2+y^2}$

→ apply def. differentiability.

$\lim_{(x,y)\to(0,0)} \dfrac{\left|\frac{x^4+y^4}{x^2+y^2}\right|}{\sqrt{x^2+y^2}} = 0$

squeeze theorem: lim = 0

∴ f diff' at $(0,0)$

(b) → $f_x = \lim_{h\to 0} \dfrac{\frac{(x+h)|y|}{\sqrt{(x+h)^2+y^2}} - \frac{x|y|}{\sqrt{x^2+y^2}}}{h} = \lim_{h\to 0} \dfrac{\frac{0}{h}}{h} = 0$

$f_y = \lim_{h\to 0} \dfrac{\frac{x|y+h|}{\sqrt{x^2+(y+h)^2}} - 0}{h} = 0$

→ $L_{(a,b)}(0,0) = \dfrac{a|b|}{\sqrt{a^2+b^2}} = 0$

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

approach along $y = mx$

$\lim_{x\to 0} \dfrac{x|mx|}{x^2+m^2x^2} = \dfrac{|m|}{1+m^2}$ depend on m ∴ DNE

not diff.

6. Let $f(x,y) = \begin{cases} \dfrac{x^3 - y^4}{x^2 + y^2} + 1 & \text{if } (x,y) \neq (0,0) \\ 1 & \text{if } (x,y) = (0,0) \end{cases}$.

Determine all points where f is differentiable.

$\to f_x = \lim\limits_{h \to 0} \dfrac{\frac{(x+h)^3}{(x+h)^2} + 1 - 1}{h} = h = 0$

$f_y = \lim\limits_{h \to 0} \dfrac{\frac{-(y+h)^4}{(y+h)^2} + 1 - 1}{h} = -h = 0$

$\to L_{(0,0)}(x,y) = 1$

$\to \lim\limits_{(x,y)\to(0,0)} \dfrac{|R_{1,(0,0)}(x,y)|}{\|(x,y)-(0,0)\|} = \lim\limits_{(x,y)\to(0,0)} \dfrac{\left|\frac{x^3-y^4}{x^2+y^2} + 1 - 1\right|}{\sqrt{x^2+y^2}} = \lim\limits_{(x,y)\to(0,0)} \dfrac{|x^3-y^4|}{|x^2+y^2|^{\frac{3}{2}}}$

approach along $y = mx$

$\lim\limits_{(x,y)\to(0,0)} \dfrac{|x^3 - m^4x^4|}{|x^2+m^2x^2|^{\frac{3}{2}}} = \lim\limits_{(x,y)\to(0,0)} \dfrac{|x^3(1-m^4)|}{|x^2(1+m^2)|^{\frac{3}{2}}} = \dfrac{1-m^4}{((1+m^2))^{\frac{3}{2}}}$

depend on m

not diff at $(0,0)$ $\therefore f$ diff at $(x,y) \neq (0,0)$

7. Let $f(x,y) = \begin{cases} \dfrac{xy^2 + y^3}{x^2 + y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$

(a) Prove that f is continuous at $(0,0)$.

(b) Determine all points where f is differentiable.

(a) $\to$ approach along $y = 0$.

$\lim\limits_{x \to 0} \dfrac{0}{x^2} = 0$

candidate limit $L = 0$

$\lim\limits_{(x,y)\to(0,0)} \left|\dfrac{xy^2+y^3}{x^2+y^2}\right| \le \lim\limits_{(0,0)} \dfrac{|y^2(x+y)|}{x^2+y^2}$ ← 先看分子是否有能提出来的

$= \lim\limits_{(0,0)} \dfrac{|x+y|\, y^2}{x^2+y^2}$

$\le \lim\limits_{(0,0)} \dfrac{|x+y|(x^2+y^2)}{x^2+y^2}$

$= \lim\limits_{(0,0)} |x+y|$

$= 0$

$\therefore$ cont at $(0,0)$

(b) f is diff at all $(x,y) \neq (0,0)$

check whether f diff at $(0,0)$

$f_x(0,0) = \lim\limits_{h \to 0} \dfrac{\frac{(x+h)\cdot 0}{(x+h)^2} - 0}{h} = 0$

$f_y(0,0) = \lim\limits_{h \to 0} \dfrac{\frac{(y+h)^3}{(y+h)^2} - 0}{h} = 1$

$\lim\limits_{(x,y)\to(0,0)} \dfrac{|R_{1,(0,0)}(x,y)|}{\|(x,y)-(0,0)\|}$

$= \lim\limits_{(x,y)\to(0,0)} \dfrac{\left|\frac{xy^2+y^3}{x^2+y^2} - (0+y)\right|}{\sqrt{x^2+y^2}}$

$= \lim\limits_{(x,y)\to(0,0)} \dfrac{\left|\frac{xy^2+\cancel{y^3}-x^2y-\cancel{y^3}}{x^2+y^2}\right|}{\sqrt{x^2+y^2}}$

$= \lim\limits_{(x,y)\to(0,0)} \dfrac{|xy(y-x)|}{(x^2+y^2)^{\frac{3}{2}}}$

$\to$ approach along $y = mx$

$\lim\limits_{x \to 0} \dfrac{|mx^2(mx-x)|}{(x^2+m^2x^2)^{\frac{3}{2}}}$

$= \lim\limits_{x \to 0} \dfrac{|m^2x^3 - mx^3|}{x^3(1+m^2)^{\frac{3}{2}}}$

$= \lim\limits_{x \to 0} \dfrac{|m^2-m|}{(1+m^2)^{\frac{3}{2}}}$

depend on m $\therefore$ DNE

8. Determine whether the functions in #4 are differentiable at $(0,a)$, $a \neq 0$.

Hint: Does $f_x(0,a)$ exist? Consider the cross-section $y = a$ to get a geometric interpretation.

(i) $f(x,y) = (xy)^{2/3}$

(ii) $f(x,y) = (xy)^{1/3}$

(iii) $f(x,y) = |x|^{\frac{1}{2}}|y|^{\frac{3}{2}}$

(iv) $f(x,y) = \begin{cases} \dfrac{x^3 + y^4}{x^2 + y^2} & \text{for } (x,y) \neq (0,0) \\ 0 & \text{for } (x,y) = (0,0) \end{cases}$

(i) $f_x(x,y) = \lim\limits_{h \to 0} \dfrac{((x+h)\cdot y)^{\frac{2}{3}} - 0}{h}$

$f_x(0,a) = \lim\limits_{h \to 0} \dfrac{h^{\frac{2}{3}} a^{\frac{2}{3}}}{h} = \lim\limits_{h \to 0} h^{-\frac{1}{3}} a^{\frac{2}{3}} = \infty$

DNE

(ii) $f_x(x,y) = \lim\limits_{h \to 0} \dfrac{((x+h)y)^{\frac{1}{3}} - (xy)^{\frac{1}{3}}}{h}$

$f_x(0,a) = \lim\limits_{h \to 0} \dfrac{h^{\frac{1}{3}} a^{\frac{1}{3}}}{h} = \lim\limits_{h \to 0} h^{-\frac{2}{3}} a^{\frac{1}{3}} = \infty$

DNE

(iii) $f_x(x,y) = \lim\limits_{h \to 0} \dfrac{|x+h|^{\frac{1}{2}}|y|^{\frac{3}{2}} - |x|^{\frac{1}{2}}|y|^{\frac{3}{2}}}{h}$

$f_x(0,a) = \lim\limits_{h \to 0} \dfrac{|h|^{\frac{1}{2}}|a|^{\frac{3}{2}}}{h} = \lim\limits_{h \to 0} \left(\dfrac{h}{\sqrt{|h|}}\right)^{-\frac{1}{2}} |a|^{\frac{3}{2}} = \infty$

DNE.

14. Prove that if f_{xx}, f_{xy}, f_{yx} and f_{yy} are continuous at (a,b), then f_x, f_y and f are continuous at (a,b).

Hint: Apply the theorems relating to differentiability.

for f_x,

$\because f_{xy}$ & f_{xx} exist

$\therefore R_{x,1,(a,b)}(x,y) = f_x(x,y) - f_x(a,b) - f_{xx}(a,b)(x-a) - f_{xy}(x,y)(y-b)$

$\boxed{\because f_{xx} \,\&\, f_{xy} \text{ cts} \quad \therefore f_x(x,y) - f_x(a,b) = f_x(x-a, y-b)}$

$\underline{f_{xx}.\ f_{xy}}$ are partial of $\underline{f_x}$.

$\underline{f_{xx}.\ f_{xy}} \downarrow$ cts $\longrightarrow$ $\underline{f_x} \downarrow$ diff' $\downarrow$ cts

THEOREM 1 **(Mean Value Theorem)**

If $f(t)$ is continuous on the closed interval $[t_1, t_2]$ and f is differentiable on the open interval (t_1, t_2), then there exists $t_0 \in (t_1, t_2)$ such that

$$f(t_2) - f(t_1) = f'(t_0)(t_2 - t_1)$$

THEOREM 2 If $\dfrac{\partial f}{\partial x}$ and $\dfrac{\partial f}{\partial y}$ are continuous at (a,b), then $f(x,y)$ is differentiable at (a,b).

9. Determine all points where f is differentiable.

(a) $f(x,y)=\begin{cases}\dfrac{x^3+y^3}{x^2+y^2} & \text{if } (x,y)\neq(0,0)\\ 0 & \text{if } (x,y)=(0,0)\end{cases}$

(b) $g(x,y)=|x|^{1/2}$

(a) $\because x^3+y^3$ & x^2+y^2 are polynomial

$\therefore f_x$ & f_y is defined everywhere except possibly at $(0,0)$

→ check whether it is diff' at $(0,0)$

$\rightarrow f_x=\dfrac{3x^2(x^2+y^2)-(x^3+y^3)(2x)}{(x^2+y^2)^2}$ $\quad f_x(0,0)=1$

$f_y=\dfrac{3y^2(x^2+y^2)-(x^3+y^3)(2y)}{(x^2+y^2)^2}$ $\quad f_y(0,0)=1$

$\rightarrow L_{(0,0)}(x,y)=0+x+y=x+y$

$\rightarrow R_{1(0,0)}(x,y)=f(x,y)-L_{(0,0)}(x,y)$

$=\dfrac{x^3+y^3}{x^2+y^2}-x-y$

$$\lim_{(0,0)}\frac{|R_{1,(0,0)}(x,y)|}{\|(x,y)-(0,0)\|}=\lim_{(0,0)}\frac{\left|\frac{x^3+y^3}{x^2+y^2}-x-y\right|}{\sqrt{x^2+y^2}}$$

$$=\lim_{(0,0)}\frac{|-xy^2-yx^2|}{(x^2+y^2)^{\frac{3}{2}}}$$

$$\lim_{(0,0)}\frac{|-xy^2-yx^2|}{(x^2+y^2)^{\frac{3}{2}}}=\lim_{x\to 0}\frac{|-m^2x^2-mx^3|}{(x^2+m^2x^2)^{\frac{3}{2}}}$$

$$=\frac{|m^2+m|}{(1+m^2)^{\frac{3}{2}}}$$

limit depend on m.

$\therefore f(x,y)$ is not diff' at $(0,0)$

(b) 遇到绝对值，先拆分成去绝对值的形式

$$g(x,y)=\begin{cases}x^{\frac{1}{2}} & x>0\\ -x^{\frac{1}{2}} & x<0\end{cases}$$

$\rightarrow f>0 \quad f_x=\frac{1}{2}x^{-\frac{1}{2}} \rightarrow \frac{1}{2\sqrt{x}}$

$f<0 \quad f_x=\cancel{-\frac{1}{2}x^{-\frac{1}{2}}}$ $\quad \frac{1}{2}(-x)^{-\frac{1}{2}}$

$f_y=0$

调[illegible]0

$\rightarrow f_x$ is defined except $x=0$

f_y is defind everywhere.

$f_x(0,y)=\lim_{h\to 0}\frac{|h|^{\frac{1}{2}}}{h}$ DNE

$\therefore g(x,y)$ is not diff at $x=0$

10. (a) Invent a function $f(x,y)$ which is continuous on $\mathbb{R}^2$ but not differentiable at $(1,2)$. Sketch the surface $z=f(x,y)$.

(b) Invent a function $f(x,y)$ which is continuous on $\mathbb{R}^2$ but not differentiable at all points of the circle $x^2+y^2=1$. Sketch the surface $z=f(x,y)$.

(a)

$f(x,y)=\dfrac{1}{2x-y}$

12. Let $f(x,y)=\begin{cases}\dfrac{x^4y^2}{x^2+y^2}+1 & \text{if } (x,y)\neq(0,0)\\ 1 & \text{if } (x,y)=(0,0).\end{cases}$

(a) Determine $f_x(0,0)$ and $f_y(0,0)$.

(b) Determine $f_x(x,y)$ for all $(x,y)\neq(0,0)$.

(c) Determine if f_x is continuous at $(0,0)$.

(a) $f_x(0,0)=\lim_{h\to 0}\dfrac{f(h,0)-1}{h}$

$=\lim_{h\to 0}\dfrac{h^4\cdot 0}{h^2+0}\cdot\dfrac{1}{h}$

$=\lim_{h\to 0}h$

$=0$

$f_y(0,0)=\lim_{h\to 0}\dfrac{f(0,h)-1}{h}$

$=\lim_{h\to 0}\dfrac{0\cdot h^2}{0\cdot h^2}\cdot\dfrac{1}{h}$

$=\infty$

(b) $f_x(x,y)=\dfrac{4x^3y^2(x^2+y^2)-x^4y^2\cdot 2x}{(x^2+y^2)^2}$

(c) → by (a) $f_x(0,0)$ exist

$\rightarrow \lim_{(x,y)\to(0,0)}\dfrac{4x^3y^2(x^2+y^2)-x^4y^2\cdot 2x}{(x^2+y^2)^2}$

$=\lim_{x\to 0}\dfrac{0}{x^4}$

$=\infty\neq 0.$

No

Chapter 6

The Chain Rule

6.1 Basic Chain Rule in Two Dimensions

Review of the Chain Rule for $\mathbf{f}(\mathbf{x}(\mathbf{t}))$

Let $T = f(x)$ be the temperature of a heated metal rod as a function of the position x. An ant runs on the rod with its position given by $x = x(t)$ as a function of time t. We want to find an expression for the rate of change of temperature with respect to time as experienced by the ant.

Observe that at any time t, the temperature at the ant's location is given by the composition of functions

$$T(t) = f(x(t))$$

Thus, the rate of change of temperature with respect to time as experienced by the ant is given by the derivative of this with respect to t. That is

$$T'(t) = f'(x(t))x'(t) \tag{6.1}$$

If we rewrite this using the Leibniz form of the Chain Rule, we get

$$\underset{\substack{\uparrow \\ T \text{ as a composite of } t}}{\frac{dT}{dt}} = \underset{\substack{\uparrow \\ T \text{ as a function of } x}}{\frac{dT}{dx}}\frac{dx}{dt} \tag{6.2}$$

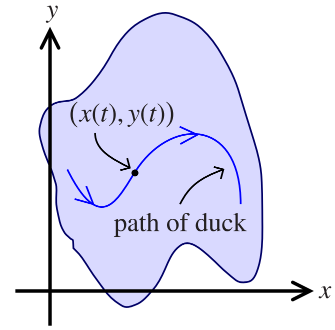
Observe that this involves an abuse of notation, since T is used in two different contexts. It is essential in what follows to understand these different ways of writing the 1-D Chain Rule.

The Chain Rule for $\mathbf{f}(\mathbf{x}(\mathbf{t}), \mathbf{y}(\mathbf{t}))$

In order to provide a physical context, suppose that the surface temperature of a pond is $T = f(x, y)$, as a function of position (x, y). A duck swims on the pond with its position given by

$$x = x(t), \quad y = y(t)$$

as a function of time t. Find an expression for the rate of change of temperature with respect to time as experienced by the duck.



The temperature experienced by the duck as a function of time t is given by the composition of functions

$$T(t) = f(x(t), y(t))$$

In a time change Δt, x and y change by

$$\Delta x = x(t + \Delta t) - x(t), \quad \Delta y = y(t + \Delta t) - y(t)$$

By the increment form of the linear approximation, the change in T corresponding to changes Δx and Δy is approximated by

$$\Delta T \approx \frac{\partial T}{\partial x}\Delta x + \frac{\partial T}{\partial y}\Delta y$$

for Δx and Δy sufficiently small. Divide by Δt, let $\Delta t \to 0$, and use the definition of the derivative to get $\frac{dT}{dt}$ on the left side of the equation. Assuming that $T(x, y)$ is differentiable at (x, y), then as Δx and $\Delta y \to 0$, the error in the linear approximation tends to zero, and so the approximation becomes increasingly accurate, leading to

$$\frac{dT}{dt} = \frac{\partial T}{\partial x}\frac{dx}{dt} + \frac{\partial T}{\partial y}\frac{dy}{dt} \tag{6.3}$$

T as a composite function of t — T as a function of x and y

This is the simplest example of the Chain Rule in two dimensions, and should be compared with equation (6.2). A precise form of equation (6.3), which avoids abuse of notation, is

$$\frac{d}{dt}f(x(t), y(t)) = f_x(x(t), y(t))x'(t) + f_y(x(t), y(t))y'(t) \tag{6.4}$$

which should be compared with equation (6.1). Alternatively, define the composite function T by

$$T(t) = f(x(t), y(t))$$

and write

$$T'(t) = f_x(x(t), y(t))x'(t) + f_y(x(t), y(t))y'(t) \tag{6.5}$$

Note that $f_x(x(t), y(t))$ is the partial derivative of the function $f(x, y)$ with respect to x, evaluated at $(x(t), y(t))$. In order to be able to apply the Chain Rule, it is important to study and understand both forms (6.3) and (6.4)/(6.5).

REMARK

The preceding "derivation" is intended to make the Chain Rule plausible, but is NOT a proof. The difficulty lies in the approximation sign $\approx$. This can be remedied by keeping track of the error in the linear approximation and leads to a proof. Note that a hypothesis on the function f, stronger than existence of the partial derivatives, is required.

THEOREM 1 **(Chain Rule)**

Let $G(t) = f(x(t), y(t))$, and let $a = x(t_0)$ and $b = y(t_0)$. If f is differentiable at (a, b) and $x'(t_0)$ and $y'(t_0)$ exist, then $G'(t_0)$ exists and is given by

$$G'(t_0) = f_x(a, b)x'(t_0) + f_y(a, b)y'(t_0)$$

Proof: By definition of the derivative,

$$G'(t_0) = \lim_{t \to t_0} \frac{G(t) - G(t_0)}{t - t_0} \tag{6.6}$$

provided that this limit exists. By definition of $G(t)$,

$$G(t) - G(t_0) = f(x(t), y(t)) - f(x(t_0), y(t_0)) \tag{6.7}$$

Since f is differentiable we can write

$$f(x, y) = f(a, b) + f_x(a, b)(x - a) + f_y(a, b)(y - b) + R_{1,(a,b)}(x, y) \tag{6.8}$$

where

$$\lim_{(x,y) \to (a,b)} \frac{|R_{1,(a,b)}(x, y)|}{\sqrt{(x - a)^2 + (y - b)^2}} = 0 \tag{6.9}$$

Since $a = x(t_0)$, $b = y(t_0)$, it follows from equations (6.7) and (6.8) that

$$\frac{G(t) - G(t_0)}{t - t_0} = f_x(a, b)\left[\frac{x(t) - x(t_0)}{t - t_0}\right] + f_y(a, b)\left[\frac{y(t) - y(t_0)}{t - t_0}\right] + \frac{R_{1,(a,b)}(x(t), y(t))}{t - t_0} \tag{6.10}$$

You can now see the Chain Rule taking shape. We have to prove that

$$\lim_{t \to t_0} \frac{|R_{1,(a,b)}(x(t), y(t))|}{|t - t_0|} = 0$$

Define $E(x, y)$ by

$$E(x, y) = \begin{cases} \dfrac{R_{1,(a,b)}(x, y)}{\sqrt{(x - a)^2 + (y - b)^2}} & \text{if } (x, y) \neq (a, b) \\ 0 & \text{if } (x, y) = (a, b) \end{cases}$$

By equation (6.9) and the definition of continuity, E is continuous at (a, b).

From the definition of E,

$$R_{1,(a,b)}(x, y) = E(x, y)\sqrt{(x-a)^2 + (y-b)^2}, \quad \text{for all } (x, y)$$

Since $a = x(t_0)$, and $b = y(t_0)$,

$$\frac{\left|R_{1,(a,b)}\big(x(t), y(t)\big)\right|}{|t-t_0|} = \left|E\big(x(t), y(t)\big)\right| \sqrt{\left[\frac{x(t)-x(t_0)}{t-t_0}\right]^2 + \left[\frac{y(t)-y(t_0)}{t-t_0}\right]^2}$$

Since $x'(t_0)$ and $y'(t_0)$ exist and the fact that E is continuous at (a, b) we get

$$\lim_{t\to t_0} \frac{\left|R_{1,(a,b)}\big(x(t), y(t)\big)\right|}{|t-t_0|} = E\big(x(t_0), y(t_0)\big)\sqrt{[x'(t_0)]^2 + [y'(t_0)]^2} = 0$$

since $E(a, b) = 0$.

It now follows from equation (6.6) and (6.10) that $G'(t_0)$ exists, and is given by the desired chain rule formula. □

REMARK

When first studying the Chain Rule you might think that hypothesis that f is differentiable could be replaced by the weaker hypothesis that $f_x(a, b)$ and $f_y(a, b)$ exist. Exercise 1 shows that this is not the case.

EXERCISE 1

With reference to the theorem, let

$$f(x, y) = (xy)^{\frac{1}{3}}, \quad x(t) = t, \quad y(t) = t^2$$

Define $G(t) = f(x(t), y(t))$ and show that $G'(0) = 1$. Further show that $f_x(0, 0) = 0$ and $f_y(0, 0) = 0$, so that the Chain Rule fails. Draw a conclusion about f at $(0, 0)$.

REMARK

In practice it is convenient to use stronger hypotheses in the Chain Rule. In particular, we usually assume that f has continuous partial derivatives at (a, b) and $x'(t)$ and $y'(t)$ are both continuous at t_0. This also allows one to obtain the stronger conclusion that $G'(t)$ is continuous at t_0. These hypotheses can usually be checked quickly, either by using the Continuity Theorems, or in more theoretical situations, by using given information.

$G(t) = (t \cdot t^2)^{\frac{1}{3}} = t$ $\quad G'(t) = 1$ $\quad G'(0) = 1$ $\quad$ 不同

$f_x(0,0) = \frac{1}{3} y^{\frac{1}{3}} x^{-\frac{2}{3}} = 0$ $\quad f_y(0,0) = 0$ $\quad G'(0) = f_x(0,0)\,x'(0) + f_y(0,0)\,y'(0) = 0$

$\therefore$ f is not diff' at $(0,0)$

EXAMPLE 1 Suppose that the temperature at position (x, y) in a pond is

$$T(x, y) = 10e^{-\frac{1}{10}(x^2+y^2)}$$

The path of a duck swimming on the pond is

$$x(t) = 2\cos t, \quad y(t) = 4\sin t$$

Find the rate of change of the pond's temperature as experienced by the duck at time $t = \dfrac{3\pi}{4}$.

Solution: Notice that the temperature along the duck's path is given by

$$T(t) = T(x(t), y(t))$$

Since T, x and y are differentiable the Chain Rule gives

$$\frac{dT}{dt} = \frac{\partial T}{\partial x}\frac{dx}{dt} + \frac{\partial T}{\partial y}\frac{dy}{dt}$$

Calculating $\dfrac{dx}{dt}$ and $\dfrac{dy}{dt}$ at $t = \dfrac{3\pi}{4}$, we obtain

$$\frac{dx}{dt}\left(\frac{3\pi}{4}\right) = -\sqrt{2}, \quad \frac{dy}{dt}\left(\frac{3\pi}{4}\right) = -2\sqrt{2}$$

At $t = \dfrac{3\pi}{4}$, the position of the duck is $(x, y) = (-\sqrt{2}, 2\sqrt{2})$. Calculate $\dfrac{\partial T}{\partial x}$ and $\dfrac{\partial T}{\partial y}$ at $(-\sqrt{2}, 2\sqrt{2})$, obtaining

$$\frac{\partial T}{\partial x}(-\sqrt{2}, 2\sqrt{2}) = \frac{2\sqrt{2}}{e}, \quad \frac{\partial T}{\partial y}(-\sqrt{2}, 2\sqrt{2}) = -\frac{4\sqrt{2}}{e}$$

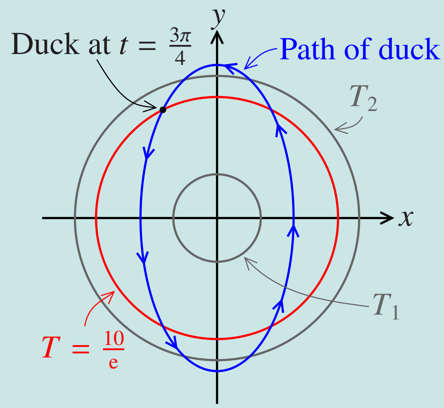
So, the Chain Rule gives

$$\frac{dT}{dt}\left(\frac{3\pi}{4}\right) = \left(\frac{2\sqrt{2}}{e}\right)(-\sqrt{2}) + \left(\frac{-4\sqrt{2}}{e}\right)(-2\sqrt{2}) = \frac{12}{e}$$

degrees/unit time.

REMARK

One can interpret the result geometrically in terms of the path of the duck and the level curves of the temperature function (the isothermal curves).
The level curves $T = \frac{10}{e}$, $T = T_1 > \frac{10}{e}$ and $T = T_2 < \frac{10}{e}$ are shown. The path of the duck is an ellipse. At time $t = \frac{3\pi}{4}$, the duck is moving from the region with $T < \frac{10}{e}$ to the region with $T = \frac{10}{e}$. Hence, we expect that $\dfrac{dT}{dt} > 0$.



EXERCISE 2

Let
$$T(t) = \ln(1 + x^2 + y^2), \quad \text{with} \quad x(t) = e^t \sin t, \quad y(t) = 2e^t \cos t$$
Calculate $\dfrac{dT}{dt}$ when $t = 0$ in two ways, firstly by substituting x and y in T, and secondly by evaluating $\dfrac{dx}{dt}(0)$, $\dfrac{dy}{dt}(0)$, $\dfrac{\partial T}{\partial x}(0, 2)$ and $\dfrac{\partial T}{\partial y}(0, 2)$, and applying the Chain Rule.

EXAMPLE 2

Define $g(t) = f(t^2 + 3, e^t)$. If $\nabla f(3, 1) = (-2, 5)$, find $g'(0)$. What condition on f will guarantee the validity of your work?

Solution: First, observe that f is a function of two variables. Say $f = f(u, v)$. Thus, we have $g(t) = f(u(t), v(t))$ where $u(t) = t^2 + 3$ and $v(t) = e^t$.

Next, to apply the Chain Rule, we require that f is differentiable.

Assuming this condition, we get

$$\begin{aligned} g'(t) &= f_x(u(t), v(t))u'(t) + f_y(u(t), v(t))v'(t) \\ &= f_x(u(t), v(t))(2t) + f_y(u(t), v(t))(e^t) \end{aligned}$$

Taking $t = 0$ gives

$$\begin{aligned} g'(0) &= f_x(u(0), v(0))(2(0)) + f_y(u(0), v(0))(e^0) \\ &= 0 + f_y(3, 1)(1) \\ &= 5 \end{aligned}$$

since $\nabla f(3, 1) = (-2, 5)$.

EXERCISE 3

Define $f(t) = g(1 + t^2, 1 - t^2)$. If $\nabla g(2, 0) = (3, 4)$, find $f'(1)$. What condition on g will guarantee the validity of your work?

EXERCISE 4

A differentiable function $f(x, y)$ is given, and $g(t)$ is defined by

$$g(t) = f(x, y)$$

where $x(t) = \cos t$ and $y(t) = \sin t$. Write out the Chain Rule for $g'(t)$. Calculate $g'\left(\frac{\pi}{3}\right)$, if $\nabla f\left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right) = \left(\sqrt{3}, 4\right)$.

The Vector Form of the Basic Chain Rule

We can use the dot product to rewrite the Chain Rule into a vector form. In particular, if we have

$$T(t) = f(x(t), y(t))$$

where $f(x, y)$, $x(t)$, and $y(t)$ are differentiable, then

$$\begin{aligned}
\frac{dT}{dt} &= \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt} \\
&= \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}\right) \cdot \left(\frac{dx}{dt}, \frac{dy}{dt}\right) \\
&= \nabla f \cdot \frac{d\mathbf{x}}{dt}
\end{aligned}$$

So, we have

$$\frac{d}{dt} f(\mathbf{x}(t)) = \nabla f(\mathbf{x}(t)) \cdot \frac{d\mathbf{x}}{dt}(t)$$

with $\mathbf{x}(t) = (x(t), y(t))$.

In this vector form, the Chain Rule holds for any differentiable function $f(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^n$, e.g. $T = f(x, y, z)$, representing temperature or some other quantity in 3-space.

EXAMPLE 3

Let the temperature at position (x, y, z) in the vicinity of the planet Mercury be given by $T = T(x, y, z)$ where T is differentiable. If the path of a spaceship is $(x(t), y(t), z(t))$, then write the Chain Rule for $\dfrac{dT}{dt}$.

Solution: We have

$$\begin{aligned}
\frac{dT}{dt} &= \nabla T(x(t), y(t), z(t)) \cdot (x'(t), y'(t), z'(t)) \\
&= T_x(x(t), y(t), z(t))x'(t) + T_y(x(t), y(t), z(t))y'(t) + T_z(x(t), y(t), z(t))z'(t)
\end{aligned}$$

EXERCISE 5

A differentiable function $f(x, y, z)$ is given and $g(t)$ is defined by

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

where $x(t) = t$, $y(t) = t^2$, and $z(t) = t^3$. Write out the Chain Rule for $g'(t)$. Find $g'(1)$ if $\nabla f(1, 1, 1) = \left(2, \frac{1}{2}, 1\right)$.

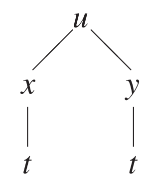
6.2 Extensions of the Basic Chain Rule

So far, we have considered composite functions formed from differentiable functions

$$u = f(x, y), \qquad \text{with} \qquad x = x(t), \quad y = y(t)$$

In this situation, the different variables are referred to as follows:

$$\begin{aligned} u &: \text{ dependent variable} \\ x, y &: \text{ intermediate variables} \\ t &: \text{ independent variable} \end{aligned}$$



The tree diagram illustrates the "chain of dependence". Observe, that our chain rule above makes sense from the point of view of rate of change. From the dependence diagram, we clearly see that the values of u are dependent on x and y which are each dependent on t. Thus, the rate of change of u should be the sum of the rate of change with respect to its x-component and with respect to its y-component. The term $\dfrac{\partial u}{\partial x}\dfrac{dx}{dt}$ calculates the rate of change of u with respect to those t's that affect u through x. Similarly $\dfrac{\partial u}{\partial y}\dfrac{dy}{dt}$ calculates the rate of change of u with respect to those t's that affect u through y.

We now discuss the case where there is more than one independent variable.

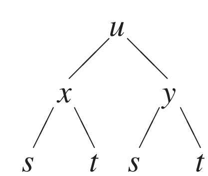
Assume $x = x(s, t)$ and $y = y(s, t)$ have first order partial derivatives at (s, t) and let

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

where f is differentiable at $(x, y) = (x(s, t), y(s, t))$. Then u is a composite function of two independent variables s and t.

Since u is a function of two variables, we want to write a chain rule for $\dfrac{\partial u}{\partial s}$ and $\dfrac{\partial u}{\partial t}$. We observe this is very similar to the case above. For $\dfrac{\partial u}{\partial s}$, the rate of change of u with respect to those s's that affect u through x is now $\dfrac{\partial u}{\partial x}\dfrac{\partial x}{\partial s}$, since x is a function of two variables. Continuing this we get

$$\begin{aligned} \frac{\partial u}{\partial s} &= \frac{\partial u}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial s} \\ \frac{\partial u}{\partial t} &= \frac{\partial u}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial u}{\partial y}\frac{\partial y}{\partial t} \end{aligned} \tag{6.11}$$



EXERCISE 1 Show that this form of the Chain Rule could also be motivated using the linear approximation. Where is the condition that $f(x, y)$ is differentiable used?

REMARKS

1. It is important to understand the difference between the various partial derivatives in equations (6.11), and to know which variable is held constant. For example

   $\dfrac{\partial u}{\partial x}$ means : regard u as the given function of x and y, and differentiate with respect to x, holding y fixed.

   $\dfrac{\partial u}{\partial s}$ means : regard u as the composite function of s and t, and differentiate with respect to s, holding t fixed.

2. Equations of the form $x = x(s, t)$, $y = y(s, t)$ can be thought of as defining a change of coordinates in 2-space.

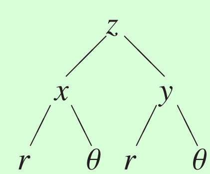
EXAMPLE 1 Let $z = f(x, y)$, where $x = r\cos\theta$, and $y = r\sin\theta$. Assuming that f is differentiable, verify that

$$\left(\frac{\partial z}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial z}{\partial \theta}\right)^2 = \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2$$

Solution: From the Chain Rule we obtain

$$\frac{\partial z}{\partial r} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial r} = \frac{\partial f}{\partial x}\cos\theta + \frac{\partial f}{\partial y}\sin\theta$$

$$\frac{\partial z}{\partial \theta} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial \theta} = \frac{\partial f}{\partial x}(-r\sin\theta) + \frac{\partial f}{\partial y}(r\cos\theta)$$



Thus, we get

$$\begin{aligned}
\left(\frac{\partial z}{\partial r}\right)^2 + \frac{1}{r^2}\left(\frac{\partial z}{\partial \theta}\right)^2 &= \left(\frac{\partial f}{\partial x}\cos\theta + \frac{\partial f}{\partial y}\sin\theta\right)^2 + \frac{1}{r^2}\left(\frac{\partial f}{\partial x}(-r\sin\theta) + \frac{\partial f}{\partial y}(r\cos\theta)\right)^2 \\
&= \left(\frac{\partial f}{\partial x}\right)^2\cos^2\theta + \left(\frac{\partial f}{\partial y}\right)^2\sin^2\theta + \left(\frac{\partial f}{\partial x}\right)^2\sin^2\theta + \left(\frac{\partial f}{\partial y}\right)^2\cos^2\theta \\
&= \left(\frac{\partial f}{\partial x}\right)^2(\cos^2\theta + \sin^2\theta) + \left(\frac{\partial f}{\partial y}\right)^2(\cos^2\theta + \sin^2\theta) \\
&= \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2
\end{aligned}$$

as required.

REMARK

In some situations (see the example to follow) it is necessary to write a more precise form of the Chain Rule (6.11), one which displays the functional dependence.

Let g denote the composite function of $f(x, y)$ and $x(s, t)$, $y(s, t)$:

$$g(s, t) = f(x(s, t), y(s, t))$$

Then, the first equation in (6.11) can be written as

$$\frac{\partial g}{\partial s}(s, t) = \frac{\partial f}{\partial x}(x(s, t), y(s, t))\frac{\partial x}{\partial s}(s, t) + \frac{\partial f}{\partial y}(x(s, t), y(s, t))\frac{\partial y}{\partial s}(s, t)$$

with a similar equation for $\dfrac{\partial g}{\partial t}(s, t)$.

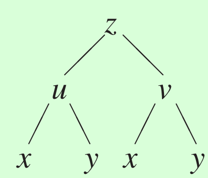
EXAMPLE 2 A differentiable function f is given with $\nabla f(2, 0) = (2, 3)$. Calculate $\dfrac{\partial g}{\partial x}(1, 1)$ where

$$g(x, y) = f(2xy, x^2 - y^2)$$

Solution: We see that f is a function of two variables. Say, $f = f(u, v)$.

Thus, we have

$$z = g(x, y) = f(u(x, y), v(x, y))$$



where $u(x, y) = 2xy$ and $v(x, y) = x^2 - y^2$.
The Chain Rule reads:

$$\begin{aligned}\frac{\partial g}{\partial x}(x, y) &= \frac{\partial f}{\partial u}(u(x, y), v(x, y))\frac{\partial u}{\partial x}(x, y) + \frac{\partial f}{\partial v}(u(x, y), v(x, y))\frac{\partial v}{\partial x}(x, y)\\ &= 2y\frac{\partial f}{\partial u}(u(x, y), v(x, y)) + 2x\frac{\partial f}{\partial v}(u(x, y), v(x, y))\end{aligned}$$

Taking $(x, y) = (1, 1)$, we obtain

$$\begin{aligned}\frac{\partial g}{\partial x}(1, 1) &= 2(1)\frac{\partial f}{\partial u}(u(1, 1), v(1, 1)) + 2(1)\frac{\partial f}{\partial v}(u(1, 1), v(1, 1))\\ &= 2\frac{\partial f}{\partial u}(2, 0) + 2\frac{\partial f}{\partial v}(2, 0)\\ &= 2(2) + 2(3) = 10\end{aligned}$$

EXERCISE 2 Referring to Example 2, calculate $\dfrac{\partial g}{\partial y}(1, 1)$.

EXERCISE 3 A function g is defined by

$$g(t) = f(h(t) + t, h(t) - t)$$

where $f(x, y)$ and $h(t)$ are both differentiable. Write the Chain Rule for $g'(t)$.

EXERCISE 4 Referring to Exercise 3, if $h(1) = 2$, $h'(1) = 3$ and $\nabla f(3, 1) = (2, -3)$, find $g'(1)$.

In the dependence diagrams in Examples 1 and 2 we see there are two paths leading from the dependent variable to the independent variable and this gives rise to a sum of two terms on the right side of the equation. Each path has two links (–), which results in each term being a product of two derivatives. Thus, we can use our dependence diagrams to find the Chain Rule for more complicated situations. In particular, to obtain the Chain Rule from a dependence diagram we have the following algorithm.

ALGORITHM

To write the Chain Rule from a dependence diagram we:

1. Take all possible paths from the differentiated variable to the differentiating variable.
2. For each link (–) in a given path, differentiate the upper variable with respect to the lower variable being careful to consider if this is a derivative or a partial derivative. Multiply all such derivatives in that path.
3. Add the products from step 2 together to complete the Chain Rule.

EXAMPLE 3 The temperature T of the water in a pond depends on position and time. Thus, we have temperature function $T = T(x, y, t)$. Find the rate of change of temperature experienced by a duck whose path is $x = x(t)$, $y = y(t)$ assuming that $T(x, y, t)$, $x(t)$, and $y(t)$ are all differentiable.

Solution: We have

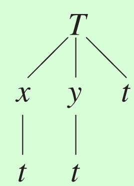
$$T = T(x, y, t), \qquad \text{where } x = x(t),\ y = y(t)$$

We draw the dependence diagram and apply the algorithm above.

The first path gives $\dfrac{\partial T}{\partial x}\dfrac{dx}{dt}$,

the second path gives $\dfrac{\partial T}{\partial y}\dfrac{dy}{dt}$,

and the third path gives $\dfrac{\partial T}{\partial t}$.



Thus, the Chain Rule is the sum of these terms. So, we have

$$\frac{dT}{dt} = \underbrace{\frac{\partial T}{\partial x}\frac{dx}{dt} + \frac{\partial T}{\partial y}\frac{dy}{dt}} + \frac{\partial T}{\partial t}$$

Rate of change of temperature with time, as experienced by duck

Contribution due to duck's movement

Due to change of temperature with time at a fixed position

It is essential to distinguish between:

$\dfrac{dT}{dt}$: the ordinary derivative of T as a composite function of t.

$\dfrac{\partial T}{\partial t}$: the partial derivative of T as the given function of x, y, t with x, y held fixed.

In order to emphasize which variables are held fixed, one can write:

$$\left(\frac{\partial T}{\partial t}\right)_{x,y}$$

In order to avoid abuse of notation, i.e. using T to denote two different functions, one can write

$$T(t) = f(x(t), y(t), t)$$

so that $T(t)$ is the function which measures the temperature at that duck's position at time t and $f(x, y, t)$ is the temperature of the water at position (x, y) at time t. Then, the Chain Rule reads

$$\frac{dT}{dt}(t) = f_x(x(t), y(t), t)x'(t) + f_y(x(t), y(t), t)y'(t) + f_t(x(t), y(t), t) \tag{6.12}$$

or more concisely

$$T'(t) = f_x x' + f_y y' + f_t$$

EXERCISE 5 Show that the Chain Rule (6.12) can also be derived by means of the increment form of the linear approximation for $f(x, y, t)$.

EXAMPLE 4 A differentiable function f is given such that $f(3, 2) = 5$ and $\nabla f(3, 2) = (4, -1)$. Let $g(t) = t^2 f(2t + 1, 3t^3 - t)$. Calculate $g'(1)$.

Solution: We see that f is a function of two variables. Say, $f = f(u, v)$. Thus, we have

$$z = g(t) = t^2 f(u(t), v(t))$$

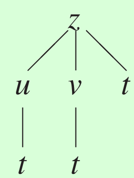
where $u(t) = 2t + 1$ and $v(t) = 3t^3 - t$.

Observe that the dependent variable z depends of the value of t and the value of $f(u, v)$. Hence, z is technically a function of three variables t, u, and v. Under composition, we get that both u and v are functions of t. Thus, we get the given dependency diagram. Using the algorithm, we find that

The first path gives $\dfrac{\partial z}{\partial u}\dfrac{du}{dt}$,

the second path gives $\dfrac{\partial z}{\partial v}\dfrac{dv}{dt}$,

and the third path gives $\left(\dfrac{\partial z}{\partial t}\right)_{u,v}$.



The Chain Rule is the sum of these terms. So,

$$g'(t) = \frac{\partial z}{\partial u}\frac{du}{dt} + \frac{\partial z}{\partial v}\frac{dv}{dt} + \left(\frac{\partial z}{\partial t}\right)_{u,v}$$

To calculate $\dfrac{\partial z}{\partial u}$, we are taking the partial derivative of z holding v and t as constants. Thus, we get

$$\frac{\partial z}{\partial u} = t^2 f_u(u, v)$$

Similarly,

$$\frac{\partial z}{\partial v} = t^2 f_v(u, v)$$

As indicated by the notation, to calculate $\left(\dfrac{\partial z}{\partial t}\right)_{u,v}$ we take the partial derivative of z holding u and v as constants. Since both u and v are considered constant, it means that $f(u, v)$ is constant. Hence, we get

$$\left(\frac{\partial z}{\partial t}\right)_{u,v} = 2t f(u, v)$$

Thus, we have that

$$g'(t) = t^2 f_u(u, v)(2) + t^2 f_v(u, v)(9t^2 - 1) + 2t f(u, v)$$

Hence,

$$\begin{aligned} g'(1) &= (1)^2 f_u(3,2)(2) + (1)^2 f_v(3,2)(9(1)^2 - 1) + 2(1)f(3,2) \\ &= 4(2) + (-1)(8) + 2(5) = 10 \end{aligned}$$

EXERCISE 6 Let f be a function of two variables such that $f(2,0) = -1$ and $\nabla f(2,0) = (2,3)$. Let $g(x,y) = xf(2xy, x^2 - y^2)$. Calculate $\dfrac{\partial g}{\partial x}(1,1)$. What assumption do you need to make about f?

EXERCISE 7 Let $u(s,t) = f(x(s,t), y(s,t), s, t)$. Write the Chain Rule for $\dfrac{\partial u}{\partial s}$, showing the functional dependence explicitly.

6.3 The Chain Rule for Second Partial Derivatives

In some situations, it is necessary to be able to calculate second derivatives of composite functions using the Chain Rule. One encounters this problem when working with partial differential equations which involve second derivatives e.g. Laplace's equation

$$u_{xx} + u_{yy} = 0$$

It also arises when working with Taylor Polynomials and in the proof of Taylor's Theorem (see Chapter 8).

Let's start with an example using functions of one variable.

EXAMPLE 1 If $z = f(x)$ where f is twice differentiable and $x = e^u$, verify that

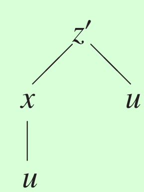
$$z''(u) = x^2 f''(x) + xf'(x)$$

Solution: Observe that by composition we have $z = z(u)$. Since $f(x)$ and $x(u)$ are differentiable the Chain Rule gives

$$z'(u) = f'(x)x'(u) = f'(x)e^u$$

Since $z'(u)$ is differentiable, we can apply the Chain Rule again to calculate $z''(u)$. Drawing the dependence diagram for $z'(u)$ and using our algorithm for calculating the Chain Rule we get

$$z''(u) = \frac{\partial z'(u)}{\partial x}\frac{dx}{du} + \left(\frac{\partial z'(u)}{\partial u}\right)_x$$



Then, we have

$$\frac{\partial z'(u)}{\partial x} = \frac{\partial f'(x)e^u}{\partial x} = f''(x)e^u$$

since we are holding u constant and taking the derivative with respect to x, and

$$\left(\frac{\partial z'(u)}{\partial u}\right)_x = \left(\frac{\partial f'(x)e^u}{\partial u}\right)_x = f'(x)e^u$$

since we are holding x constant and taking the derivative with respect to u. Finally, $\frac{dx}{du} = e^u$ and so we get

$$z''(u) = (f''(x)e^u)(e^u) + f'(x)e^u = x^2 f''(x) + xf'(x) \tag{6.13}$$

REMARK

Observe, if we had substituted in $x = e^u$ at the beginning, we would get

$$z'(u) = f'(e^u)e^u$$

Hence, taking the derivative with respect to u we would get

$$\begin{aligned} z''(u) &= \frac{d}{du}(f'(e^u))e^u + f'(e^u)\frac{d}{du}(e^u) && \text{by the Product Rule} \\ &= \left(f''(e^u)\frac{d}{du}(e^u)\right)e^u + f'(e^u)e^u && \text{by the Chain Rule} \\ &= (f''(e^u)e^u)e^u + f'(e^u)e^u \end{aligned}$$

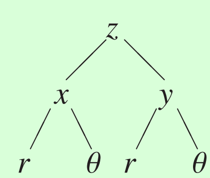
which matches (6.13). Thus, we see that our dependence diagram algorithm not only calculates the necessary Chain Rules, but also includes the necessary Product Rules.

EXAMPLE 2 Let $z = f(x, y)$ with $x = r\cos\theta$ and $y = r\sin\theta$. Verify that

$$\frac{\partial^2 z}{\partial r^2} + \frac{1}{r}\frac{\partial z}{\partial r} + \frac{1}{r^2}\frac{\partial^2 z}{\partial \theta^2} = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

What assumptions do you need to make about f?

Solution: Assuming that f is differentiable the Chain Rule gives



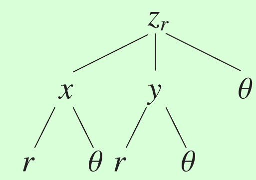
$$\frac{\partial z}{\partial r} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial r} = \frac{\partial f}{\partial x}\cos\theta + \frac{\partial f}{\partial y}\sin\theta \tag{6.14}$$

In order to calculate $\frac{\partial^2 z}{\partial r^2}$, we have to use the Chain Rule to differentiate this equation with respect to r, keeping θ constant.

To draw the dependence diagram, we first write (6.14) more precisely showing the functional dependence. It is

$$z_r(r,\theta) = f_x(x,y)\cos\theta + f_y(x,y)\sin\theta$$

So, we see that z_r is dependent on x, y and θ where, by composition, x and y are both dependent on r and θ. Thus, we get the dependence diagram to the right.



Since we will be taking partial derivatives of f_x and f_y, to apply the Chain Rule, we now need to assume that f_x and f_y are differentiable. We get

$$\frac{\partial^2 z}{\partial r^2} = \frac{\partial z_r}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial z_r}{\partial y}\frac{\partial y}{\partial r} \tag{6.15}$$

Then,

$$\begin{aligned}\frac{\partial z_r}{\partial x} &= \frac{\partial(f_x\cos\theta + f_y\sin\theta)}{\partial x}\\ &= \frac{\partial f_x}{\partial x}\cos\theta + \frac{\partial f_y}{\partial x}\sin\theta \qquad \text{since we are holding } \theta \text{ constant}\\ &= f_{xx}\cos\theta + f_{yx}\sin\theta\end{aligned}$$

Since we assumed that f_x and f_y are differentiable, we find that

$$\frac{\partial z_r}{\partial y} = f_{xy}\cos\theta + f_{yy}\sin\theta$$

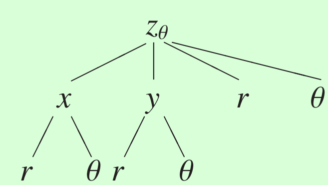
Putting these into (6.15) and computing $\dfrac{\partial x}{\partial r}$ and $\dfrac{\partial y}{\partial r}$ we find that

$$\begin{aligned}\frac{\partial^2 z}{\partial r^2} &= (f_{xx}\cos\theta + f_{yx}\sin\theta)\cos\theta + (f_{xy}\cos\theta + f_{yy}\sin\theta)\sin\theta\\ &= f_{xx}\cos^2\theta + f_{yx}\sin\theta\cos\theta + f_{xy}\cos\theta\sin\theta + f_{yy}\sin^2\theta \qquad (6.16)\end{aligned}$$

We now repeat this process to find $z_{\theta\theta}$. We have

$$\begin{aligned}z_\theta(r,\theta) &= f_x(x,y)x_\theta(r,\theta) + f_y(x,y)y_\theta(r,\theta)\\ &= f_x(x,y)(-r\sin\theta) + f_y(x,y)(r\cos\theta)\end{aligned}$$

Thus, we get the dependence diagram to the right



Assuming that f_x and f_y are differentiable the Chain Rule for $z_{\theta\theta}$ becomes

$$\frac{\partial^2 z}{\partial\theta^2} = \frac{\partial z_\theta}{\partial x}\frac{\partial x}{\partial\theta} + \frac{\partial z_\theta}{\partial y}\frac{\partial y}{\partial\theta} + \left(\frac{\partial z_\theta}{\partial\theta}\right)_{x,y,r} \tag{6.17}$$

We find that

$$\begin{aligned}
\frac{\partial z_\theta}{\partial x} &= \frac{\partial(f_x \cdot (-r\sin\theta) + f_y \cdot (r\cos\theta))}{\partial x} \\
&= \frac{\partial f_x}{\partial x} \cdot (-r\sin\theta) + \frac{\partial f_y}{\partial x} \cdot (r\cos\theta) \qquad \text{since } r, \theta \text{ are held constant} \\
&= -f_{xx} \cdot r\sin\theta + f_{yx} \cdot r\cos\theta \\
\frac{\partial z_\theta}{\partial y} &= \frac{\partial f_x}{\partial y} \cdot (-r\sin\theta) + \frac{\partial f_y}{\partial y} \cdot (r\cos\theta) \qquad \text{since } r, \theta \text{ are held constant} \\
&= -f_{xy} \cdot r\sin\theta + f_{yy} \cdot r\cos\theta \\
\frac{\partial z_\theta}{\partial \theta} &= \frac{\partial(f_x \cdot (-r\sin\theta) + f_y \cdot (r\cos\theta))}{\partial \theta} \\
&= -f_x \cdot r\frac{\partial \sin\theta}{\partial\theta} + f_y \cdot r\frac{\partial\cos\theta}{\partial\theta} \qquad \text{since } x, y, r \text{ are held constant} \\
&= -f_x \cdot r\cos\theta - f_y \cdot r\sin\theta
\end{aligned}$$

Putting these into (6.17) we get

$$\begin{aligned}
f_{\theta\theta} &= (-f_{xx} r\sin\theta + f_{yx} r\cos\theta)(-r\sin\theta) + (-f_{xy} r\sin\theta + f_{yy} r\cos\theta)(r\cos\theta) \\
&\quad + (-f_x r\cos\theta - f_y r\sin\theta) \\
&= f_{xx} r^2\sin^2\theta - f_{yx} r^2\cos\theta\sin\theta - f_{xy} r^2\sin\theta\cos\theta + f_{yy} r^2\cos^2\theta \\
&\quad + (-f_x r\cos\theta - f_y r\sin\theta) \qquad (6.18)
\end{aligned}$$

Using (6.14), (6.16) and (6.18) we get

$$\begin{aligned}
\frac{\partial^2 z}{\partial r^2} + \frac{1}{r}\frac{\partial z}{\partial r} + \frac{1}{r^2}\frac{\partial^2 z}{\partial \theta^2} &= (f_{xx}\cos^2\theta + f_{yx}\sin\theta\cos\theta + f_{xy}\cos\theta\sin\theta + f_{yy}\sin^2\theta) \\
&\quad + \frac{1}{r}(f_x\cos\theta + f_y\sin\theta) + \frac{1}{r^2}(f_{xx} r^2\sin^2\theta - f_{yx} r^2\cos\theta\sin\theta \\
&\quad - f_{xy} r^2\sin\theta\cos\theta + f_{yy} r^2\cos^2\theta + -f_x r\cos\theta - f_y r\sin\theta) \\
&= f_{xx}(\cos^2\theta + \sin^2\theta) + f_{yy}(\sin^2\theta + \cos^2\theta) \\
&= f_{xx} + f_{yy}
\end{aligned}$$

as required.

EXERCISE 1 Let $g(u, v)$ be a function, and let f be defined by

$$f(x) = g(x, 2x)$$

Verify that

$$f''(x) = g_{uu} + 4g_{uv} + 4g_{vv}$$

What assumption on g will ensure that your calculation is valid?

EXERCISE 2 A function $g(t)$ is given, and f is defined by

$$f(x, y) = g(xy)$$

Verify that

$$x^2 f_{xx} = y^2 f_{yy}$$

What assumption on g will ensure that your calculation is valid?

EXERCISE 3 Let $f(x, y) \in C^2$ and define g by

$$g(s) = f(a + hs, b + ks)$$

where (a, b) and (h, k) are regarded as fixed. Verify that

$$g'(s) = f_x(a + hs, b + ks)h + f_y(a + hs, b + ks)k$$
$$g''(s) = f_{xx}(a + hs, b + ks)h^2 + 2f_{xy}(a + hs, b + ks)hk + f_{yy}(a + hs, b + ks)k^2$$

Chapter 6 Problem Set

1. Let $w = x^2y + xy^3, x = 3t + 5, y = 2t^2 - 10$. Use the Chain Rule to calculate $\dfrac{dw}{dt}$ when $t = -2$.
2. (a) State the Chain Rule for a composite function $g(t) = f(x(t), y(t))$, clearly indicating the hypotheses and the conclusion.

   (b) Given a function of two variables f, let the single-variable function g be defined by

   $$g(t) = f(e^t \cos t, e^t \sin t)$$

   If $\nabla f(1, 0) = (8, -4)$, find $g'(0)$. What hypotheses must f satisfy?
3. Suppose that $f(x, y, z)$ is given, and that

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

   If $\nabla f(1, 1, 1) = (5, -3, -4)$, find $g'(1)$. What hypothesis must f satisfy?
4. Suppose that $f(x, y)$ is given, and that g is defined by

   $$g(s, t) = f(st, s^2 - t^2)$$

   If $\nabla f(2, -3) = (4, 3)$, find $\nabla g(1, 2)$. What hypothesis must f satisfy?
5. Write the Chain Rule for the indicated derivatives of the composite functions, assuming that the various functions are differentiable:

   (a) If $w = f(x, y, z)$, and $x = x(s, t), y = y(s, t)$, $z = z(s, t)$, find $\dfrac{\partial w}{\partial t}$.

   (b) If $z = f(x, y)$, and $y = g(x)$, find $\dfrac{dz}{dx}$.

   (c) If $z = f(x, y)$, and $y = g(x), x = h(u, v)$, find $\dfrac{\partial z}{\partial u}$.

   (d) If $w = f(x, y, z)$, and $y = g(x, z), z = h(x)$, find $\dfrac{dw}{dx}$.

   (e) If $w = F(p, q, r, s)$, and $r = f(p, q), s = g(p, q)$, find $\left(\dfrac{\partial w}{\partial p}\right)_{q=const}$.
6. For some constant θ define

   $$u(x, y) = f(x \cos\theta + y \sin\theta, -x \sin\theta + y \cos\theta)$$

   Express $u_{xx}+u_{yy}$, $u_{xx}-u_{yy}$ and u_{xy} in terms of the second partial derivatives of f. Use double angle trigonometric identities to simplify.

7. In the following questions, state the assumption that you make about f.

   (a) If $F(x,y) = yf(x^2 - y^2)$, show that
   $$y\frac{\partial F(x,y)}{\partial x} + x\frac{\partial F(x,y)}{\partial y} = \frac{x}{y}F(x,y)$$

   (b) If $u = x^3 f\left(\frac{y}{x}, \frac{z}{x}\right)$, show that
   $$x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} + z\frac{\partial u}{\partial z} = 3u$$

   (c) If $F(x,y,z) = f\left(\frac{y-z}{x}, \frac{z-x}{y}, \frac{x-y}{z}\right)$, show that $x\frac{\partial F}{\partial x} + y\frac{\partial F}{\partial y} + z\frac{\partial F}{\partial z} = 0$.

8. Assume that $f(x,y)$ has continuous second partial derivatives. Let $x = r\cos\theta$ and $y = r\sin\theta$. Show that
$$\frac{\partial^2 f}{\partial r^2} + \frac{1}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2} = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

9. Recall for a simple electrical circuit that Ohm's Law states that $V = IR$ where V is the voltage, I is the current and R is the resistance. Find the rate of change of the current when $R = 400\ \Omega$, $I = 0.08\ A$, $dV/dt = -0.01$ V/s, and $dR/dt = 0.03\ \Omega$/s. What is the significance of the sign of your answer?

10. The path of a spacecraft is given by $(x,y,z) = (e^{2t}\cos t, e^{2t}\sin t, 2t+1)$ where t denotes time. The temperature at position (x,y,z) is given by a function $u(x,y,z)$, and the temperature gradient at (1,0,1) is $\nabla u(1,0,1) = \left(\frac{1}{5}, -\frac{1}{3}, -\frac{1}{4}\right)$.

    (a) Find the velocity of the spacecraft at time t.

    (b) Find the rate of change of temperature experienced by the spacecraft at time $t = 0$.

11. A proctologist is walking around the exam room. His position is given by $(x(t), y(t)) = (\cos t, \sin t)$. At position (x,y) his cellphone gets a signal strength of $F(x,y) = e^x y^2$. Using the Chain Rule, find the rate of change of the signal strength with respect to time at $t = \pi/2$.

12. A particle travels along the path $(x,y) = (t^2 - t, e^{3t})$ in a plane where the temperature at position (x,y) and time t is given by $T(x,y,t) = 2x^2 y\sin t$. Calculate the rate of change of temperature along the particle's path with respect to time at any time t.

13. Show that if f and g are twice differentiable functions, then $u(x,t) = f(x-at) + g(x+at)$ is a solution of the wave equation: $u_{tt} = a^2 u_{xx}$.

14. Let f be a function of two variables and define $g(x,y) = f(\sin y, \cos x)$. Find g_{xx} and g_{yy}. State any assumptions you needed to make.

15. Let $g(u,v) = f(u^2 - v^2, 2uv)$. Express $(g_u)^2 + (g_v)^2$ and $g_{uu} + g_{vv}$ in terms of the partial derivatives of f. What hypothesis must f satisfy?

16. If $u = f(x + g(y))$, where f and g have a continuous second derivative, show that $u_x u_{xy} = u_y u_{xx}$.

17. A function $g(u)$ with continuous second derivative is given, and f is defined by $f(x,y) = g\left(\frac{x}{y}\right)$, for $y \neq 0$. Calculate $\frac{\partial^2 f}{\partial x \partial y}$ and $\frac{\partial^2 f}{\partial y \partial x}$ and verify that they are equal.

18. Let $f(x,y) = (xy)^{1/3}, p(t) = t, q(t) = t^2$, and consider the composite function H defined by $H(t) = f(p(t), q(t))$. Show that the Chain Rule for $H(t)$ is not satisfied at $t = 0$. What conclusion can you draw about f at $(0,0)$?

19. Let $f(x,y,z)$ and $g(x,y,z)$ have continuous partial derivatives. Prove that
$$\nabla(fg) = f\nabla g + g\nabla f$$

20. (a) Let $F(t) = f(a+th, b+tk)$, where the two-variable function f has continuous second partial derivatives, and a, b, h, k are constants. Show that
    $$F''(t) = h^2 f_{11} + 2hkf_{12} + k^2 f_{22},$$
    where f_{11}, f_{12} and f_{22} are evaluated at $(a+th, b+tk)$.

    (b) Can you generalize (a) to give a formula for $F'''(t)$?

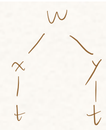
21. Functions $f(x,y,z)$ which satisfy Laplace's equation $f_{xx} + f_{yy} + f_{zz} = 0$ are of interest in theoretical physics.

    (a) Suppose that the single-variable function g has a continuous second derivative, and $f(x,y,z) = g\left(\frac{1}{r}\right)$, where $r = \sqrt{x^2 + y^2 + z^2} > 0$. Show that
    $$f_{xx} + f_{yy} + f_{zz} = \frac{1}{r^4}g''\left(\frac{1}{r}\right), \text{for } r > 0$$

    (b) Give a function f, other than a linear function, which satisfies Laplace's equation.

22. * Let the three-variable function f be differentiable and satisfy $f(t\mathbf{x}) = t^p f(\mathbf{x})$, for all $\mathbf{x} \in \mathbb{R}^3$ and $t \in \mathbb{R}$, where p is constant. Prove that
$$\mathbf{x} \cdot \nabla f(\mathbf{x}) = pf(\mathbf{x}) \quad \text{for all } \mathbf{x} \in \mathbb{R}^3$$

1. Let $w = x^2y + xy^3, x = 3t + 5, y = 2t^2 - 10$. Use the Chain Rule to calculate $\dfrac{dw}{dt}$ when $t = -2$.



$\frac{dw}{dt} = \frac{dw}{dx}\cdot\frac{dx}{dt} + \frac{dw}{dy}\cdot\frac{dy}{dt}$

$= (2xy+y^3)\cdot 3 + (x^2+3xy^2)\cdot 4t.$

$t=-2. \quad x=-1 \quad y=-2.$

$\frac{dw}{dt} = (2\cdot(-1)\cdot(-2)-8)\cdot 3 + (1+3(-1)\cdot 4)\cdot(-8)$

$= (4-8)\cdot 3 + (1-12)(-8)$

$= 76$

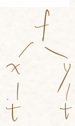
2. (a) State the Chain Rule for a composite function $g(t) = f(x(t), y(t))$, clearly indicating the hypotheses and the conclusion.

(b) Given a function of two variables f, let the single-variable function g be defined by

$$g(t) = f(\underbrace{e^t\cos t}_{x}, \underbrace{e^t\sin t}_{y})$$

If $\nabla f(1, 0) = (8, -4)$, find $g'(0)$. What hypotheses must f satisfy?

(a) $\frac{dg}{dt} = \frac{\partial f}{\partial x}\cdot\frac{dx}{dt} + \frac{\partial f}{\partial y}\cdot\frac{dy}{dt}$



条件: x & y diff at t.

(b) $g'(0) = \frac{\partial f}{\partial t}$

$= \nabla f(x(t), y(t))\cdot(x'(t), y'(t))$

$= \nabla f(x(t), y(t))\cdot(e^t\cos t - e^t\sin t, e^t\sin t + e^t\cos t)$

$= \nabla f(1,0)\cdot(1,1)$

$= (8,-4)(1,1)$

$= 4$

hypothesis: x, y diff at t.

3. Suppose that $f(x, y, z)$ is given, and that

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

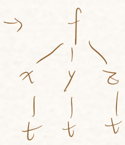
If $\nabla f(1, 1, 1) = (5, -3, -4)$, find $g'(1)$. What hypothesis must f satisfy?

→ let $f = f(x(t), y(t), z(t))$
$\qquad\quad t \qquad t^2 \qquad t^3$

→ Assume f is diff' at $(1,1,1)$

$x'(t) = 1 \qquad y'(t) = 2t \qquad z'(t) = 3t^2$

→



$g'(1) = \nabla f(1,1,1)\cdot(x'(t), y'(t), z'(t))$

$= (5,-3,-4)\cdot(1,2,3)$

$= -13$

4. Suppose that $f(x, y)$ is given, and that g is defined by

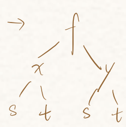
$$g(s, t) = f(st, s^2 - t^2)$$

If $\nabla f(2, -3) = (4, 3)$, find $\nabla g(1, 2)$. What hypothesis must f satisfy?

→ let $f = f(x(s,t), y(s,t))$

→ Assume f is diff' at $(2,-3)$

→



$g(1,2) = f(2,-3)$

$\nabla g(s,t) = \left(\frac{\partial g}{\partial s}, \frac{\partial g}{\partial t}\right)$

$= \left(\frac{\partial f}{\partial x}\cdot\frac{\partial x}{\partial s} + \frac{\partial f}{\partial y}\cdot\frac{\partial y}{\partial s}, \frac{\partial f}{\partial x}\cdot\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\cdot\frac{\partial y}{\partial t}\right)$

$= ((2,-3)\cdot(t,2s), (2,-3)\cdot(s,-2t))$

$= ((2,-3)\cdot(2,4), (2,-3)(1,-4))$

$= (-8, 14)$

5. Write the Chain Rule for the indicated derivatives of the composite functions, assuming that the various functions are differentiable:

   (a) If $w = f(x, y, z)$, and $x = x(s, t)$, $y = y(s, t)$, $z = z(s, t)$, find $\dfrac{\partial w}{\partial t}$.

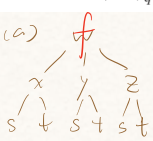
   (b) If $z = f(x, y)$, and $y = g(x)$, find $\dfrac{dz}{dx}$.

   (c) If $z = f(x, y)$, and $y = g(x)$, $x = h(u, v)$, find $\dfrac{\partial z}{\partial u}$.

   (d) If $w = f(x, y, z)$, and $y = g(x, z)$, $z = h(x)$, find $\dfrac{dw}{dx}$.

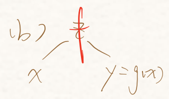
   (e) If $w = F(p, q, r, s)$, and $r = f(p, q)$, $s = g(p, q)$, find $\left(\dfrac{\partial w}{\partial p}\right)_{q=const}$.

(a)



$$\frac{\partial w}{\partial t} = \frac{\partial f}{\partial x}\cdot\frac{\partial x}{\partial t} + \frac{\partial f}{\partial y}\cdot\frac{\partial y}{\partial t} + \frac{\partial f}{\partial z}\cdot\frac{\partial z}{\partial t}.$$

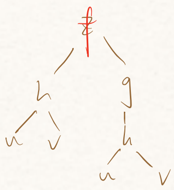
(b)



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

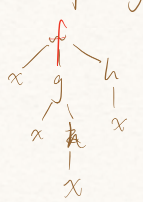
$$\frac{dz}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial g}\cdot\frac{\partial g}{\partial x}$$

(c) $z = f(h(u,v), g(h(u,v)))$.



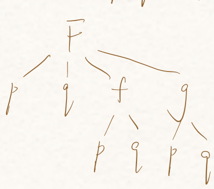
$$\frac{\partial z}{\partial u} = \frac{df}{dh}\cdot\frac{dh}{du} + \frac{df}{dg}\cdot\frac{dg}{dh}\cdot\frac{dh}{du}.$$

(d) $w = f(x, g(x, h(x)), h(x))$



$$\frac{dw}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial g}\cdot\frac{\partial g}{\partial h}\cdot\frac{\partial h}{\partial x} + \frac{\partial f}{\partial h}\cdot\frac{\partial h}{\partial x}$$

(e) $w = F(p, q, f(p,q), g(p,q))$



$$\frac{\partial w}{\partial p} = \frac{\partial F}{\partial p} + \frac{\partial F}{\partial f}\cdot\frac{\partial f}{\partial p} + \frac{\partial F}{\partial g}\cdot\frac{\partial g}{\partial p}$$

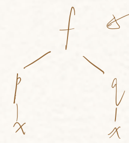
6. For some constant θ define

$$u(x, y) = f(\underbrace{x\cos\theta + y\sin\theta}_{p}, \underbrace{-x\sin\theta + y\cos\theta}_{q})$$

$\frac{\partial p}{\partial x} = \cos\theta$

Express $u_{xx}+u_{yy}$, $u_{xx}-u_{yy}$ and u_{xy} in terms of the second partial derivatives of f. Use double angle trigonometric identities to simplify.

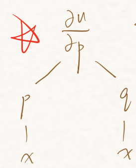
let $p = x\cos\theta + y\sin\theta$, $q = x\sin\theta + y\cos\theta$



$$\frac{\partial u}{\partial x} = \frac{\partial f}{\partial p}\cdot\frac{\partial p}{\partial x} + \frac{\partial f}{\partial q}\cdot\frac{\partial q}{\partial x}.$$

$$\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}\right) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial p}\cdot\frac{\partial p}{\partial x} + \frac{\partial f}{\partial q}\cdot\frac{\partial q}{\partial x}\right)$$

$$= \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial p}\cdot\frac{\partial p}{\partial x}\right) + \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial q}\cdot\frac{\partial q}{\partial x}\right)$$



$$= \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial p}\right)\cdot\frac{\partial p}{\partial x} + \frac{\partial f}{\partial p}\cdot\frac{\partial}{\partial x}\left(\frac{\partial p}{\partial x}\right)$$

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial q}\right)\frac{\partial q}{\partial x} + \frac{\partial f}{\partial q}\cdot\frac{\partial}{\partial x}\left(\frac{\partial q}{\partial x}\right)$$

product rule

$$= \left(\frac{\partial \frac{\partial f}{\partial p}}{\partial p}\cdot\frac{\partial p}{\partial x} + \frac{\partial \frac{\partial f}{\partial p}}{\partial q}\cdot\frac{\partial q}{\partial x}\right)\cdot\frac{\partial p}{\partial x}$$

$$+ \frac{\partial f}{\partial p}\cdot\frac{\partial}{\partial x}\left(\frac{\partial p}{\partial x}\right)$$

$$\frac{\partial}{\partial p}\left(\frac{\partial f}{\partial p}\right)$$

$$+\left(\frac{\partial \frac{\partial f}{\partial q}}{\partial p}\cdot\frac{\partial p}{dx} + \frac{\partial \frac{\partial f}{\partial q}}{\partial q}\cdot\frac{\partial q}{\partial x}\right)\cdot\frac{\partial q}{\partial x}$$

$$+ \frac{\partial f}{\partial q}\cdot\frac{\partial}{\partial x}\left(\frac{\partial q}{\partial x}\right)$$

一样.

Setting $u(x,y) = f(x\cos\theta + y\sin\theta, -x\sin\theta + y\cos\theta$, $w(x,y) = x\cos\theta + y\sin\theta$, and $v(x,y) = -x\sin\theta + y\cos\theta$ we have

$$u_{xx} = f_{ww}w_xw_x + f_{wv}w_xv_x + f_{vw}w_xv_x + f_{vv}v_xv_x + f_ww_{xx} + f_vv_{xx}$$

and

$$u_{xy} = f_{ww}w_xw_y + f_{wv}w_xv_y + f_{vw}w_yv_x + f_{vv}v_xv_y + f_ww_{xy} + f_vv_{xy}$$

7. In the following questions, state the assumption that you make about f.

(a) If $F(x, y) = yf(\underbrace{x^2 - y^2}_{z})$, show that

$$y\frac{\partial F(x, y)}{\partial x} + x\frac{\partial F(x, y)}{\partial y} = \frac{x}{y}F(x, y)$$

(b) If $u = x^3 f\left(\frac{y}{x}, \frac{z}{x}\right)$, show that

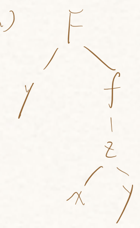
$w \to \frac{y}{x}$, $\frac{z}{x} \to v$

$$x\frac{\partial u}{\partial x} + y\frac{\partial u}{\partial y} + z\frac{\partial u}{\partial z} = 3u$$

(c) If $F(x, y, z) = f\left(\frac{y - z}{x}, \frac{z - x}{y}, \frac{x - y}{z}\right)$, show that

$$x\frac{\partial F}{\partial x} + y\frac{\partial F}{\partial y} + z\frac{\partial F}{\partial z} = 0.$$

(a)



$$\frac{\partial F(x,y)}{\partial x} = \frac{\partial F}{\partial f}\cdot\frac{df}{dz}\cdot\frac{\partial z}{\partial x}$$

$$= \cancel{\frac{\partial F}{\partial f}}\cdot f'(x^2-y^2)\cdot 2x$$

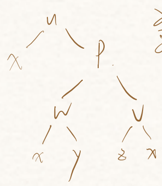
$$\frac{\partial F(x,y)}{\partial y} = \frac{\partial F}{\partial f}\cdot\frac{df}{dz}\cdot\frac{\partial z}{dy} + \frac{\partial F}{\partial y} \qquad f(x^2-y^2)$$

$$= \cancel{\frac{\partial F}{\partial f}}\cdot f'(x^2-y^2)(-2y) + \cancel{\frac{\partial F}{\partial y}}$$

$$LHS = y\left(\frac{\partial F}{\partial f}\cdot f'(x^2-y^2)\cdot 2x\right) + x\left(\frac{\partial F}{\partial f}\cdot f'(x^2-y^2)(-2y) + \frac{\partial F}{\partial y}\right)$$

$$= x\frac{\partial F}{\partial y} \qquad = xf(x^2-y^2) = \frac{x}{y}F(x,y)$$

(b) let $p = f(w, v)$



$$\frac{\partial u}{\partial x} = 3x^2 f(w,v) + x^3 f'(w,v)$$

$$= 3x^2 f + x^3\left(\frac{\partial f}{\partial w}\cdot\frac{\partial w}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x}\right)$$

$$= 3x^2 f(w,v) + x^3\frac{\partial f}{\partial w}(-yx^{-2}) + x^3\cdot\frac{\partial f}{\partial v}(-zx^{-2})$$

$$x\frac{\partial u}{\partial x} = 3x^3 f(w\,v) - yx^2\frac{\partial f}{\partial w} - zx^2\frac{\partial f}{\partial v}$$

$$y\frac{\partial u}{\partial y} = x^3 f'(w,v) = x^3\cdot\frac{\partial f}{\partial w}\cdot\frac{1}{x}\cdot y = yx^2\frac{\partial f}{\partial w}$$

$$z\frac{\partial u}{\partial z} = x^3 f'(w,v) = x^3\cdot\frac{\partial f}{\partial v}\cdot\frac{1}{x}\cdot z = zx^2\frac{\partial f}{\partial v}$$

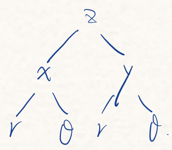
QED

8. Assume that $f(x, y)$ has continuous second partial derivatives. Let $x = r\cos\theta$ and $y = r\sin\theta$. Show that

$$\frac{\partial^2 f}{\partial r^2} + \frac{1}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2} = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

Let $z = f(x, y)$

→ first - partial derivative



$$\frac{\partial z}{\partial r} = \frac{\partial f}{\partial x}\cdot\frac{\partial x}{\partial r} + \frac{\partial f}{\partial y}\cdot\frac{\partial y}{\partial r}$$

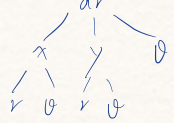
$$= \frac{\partial f}{\partial x}\cos\theta + \frac{\partial f}{\partial y}\sin\theta$$

$$\frac{\partial z}{\partial \theta} = \frac{\partial f}{\partial x}\cdot\frac{\partial x}{\partial \theta} + \frac{\partial z}{\partial y}\cdot\frac{\partial y}{\partial \theta}$$

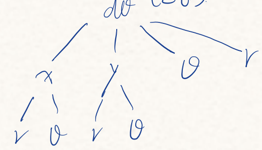
$$= \frac{\partial f}{\partial x}(-r\sin\theta) + \frac{\partial f}{\partial y}(r\cos\theta)$$

→ second - partial derivative

$\frac{dz}{dr}$ (z_r)



$\frac{dz}{d\theta}$ (z_θ)



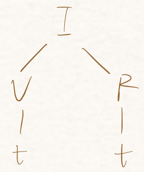
By the Chain Rule, we get

$$z_{rr} = \frac{\partial^2 z}{\partial r^2} = \frac{\partial z_r}{\partial x}\frac{\partial x}{\partial r} + \frac{\partial z_r}{\partial y}\frac{\partial y}{\partial r} \text{ and } z_{\theta\theta} = \frac{\partial^2 z}{\partial \theta^2} = \frac{\partial z_\theta}{\partial x}\frac{\partial x}{\partial \theta} + \frac{\partial z_\theta}{\partial y}\frac{\partial y}{\partial \theta} + \frac{\partial z_\theta}{\partial \theta}$$

- $\frac{\partial z_r}{\partial x} = \frac{\partial}{\partial x}(f_x\cos\theta + f_y\sin\theta) = f_{xx}\cos\theta + f_{yx}\sin\theta$
- $\frac{\partial z_r}{\partial y} = \frac{\partial}{\partial y}(f_x\cos\theta + f_y\sin\theta) = f_{xy}\cos\theta + f_{yy}\sin\theta$
- $\frac{\partial z_\theta}{\partial x} = \frac{\partial}{\partial x}\big(f_x(-r\sin\theta) + f_y(r\cos\theta)\big) = -f_{xx}(r\sin\theta) + f_{yx}(r\cos\theta)$
- $\frac{\partial z_\theta}{\partial y} = \frac{\partial}{\partial y}\big(f_x(-r\sin\theta) + f_y(r\cos\theta)\big) = -f_{xy}(r\sin\theta) + f_{yy}(r\cos\theta)$
- $\frac{\partial z_\theta}{\partial \theta} = \frac{\partial}{\partial \theta}\big(f_x(-r\sin\theta) + f_y(r\cos\theta)\big) = -f_x(r\cos\theta) - f_y(r\sin\theta)$

9. Recall for a simple electrical circuit that Ohm's Law states that $V = IR$ where V is the voltage, I is the current and R is the resistance. Find the rate of change of the current when $R = 400\ \Omega$, $I = 0.08\ A$, $dV/dt = -0.01$ V/s, and $dR/dt = 0.03\ \Omega$/s. What is the significance of the sign of your answer?

$I = \frac{V}{R}$ $\quad I = VR^{-1}$ $\quad \frac{dI}{dR} = -VR^{-2}$



$$\frac{\partial I}{\partial t} = \frac{\partial I}{\partial V}\cdot\frac{dV}{dt} + \frac{\partial I}{\partial R}\cdot\frac{dR}{dt}$$
$$= \frac{1}{R}\cdot\frac{dV}{dt} + (-VR^{-2})\,\frac{dR}{dt}$$
$$= \frac{1}{400}\cdot(-0.01) - \frac{0.08}{400}\cdot(0.03)$$
$$= -3.1\times10^{-5}\ A/s$$

10. The path of a spacecraft is given by $(x, y, z) = (e^{2t}\cos t, e^{2t}\sin t, 2t + 1)$ where t denotes time. The temperature at position (x, y, z) is given by a function $u(x, y, z)$, and the temperature gradient at (1,0,1) is $\nabla u(1, 0, 1) = \left(\frac{1}{5}, -\frac{1}{3}, -\frac{1}{4}\right)$.

(a) Find the velocity of the spacecraft at time t.

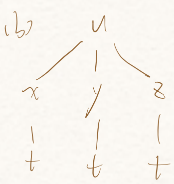
(b) Find the rate of change of temperature experienced by the spacecraft at time $t = 0$.

(a) $(x, y, z) = (f(t), g(t), h(t))$

$$\frac{dx}{dt} = 2e^{2t}\cos t - e^{2t}\sin t$$
$$\frac{dy}{dt} = 2e^{2t}\sin t + e^{2t}\cos t$$
$$\frac{dz}{dt} = 2$$

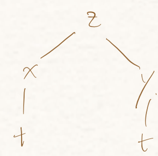
分别找出 $\frac{dx}{dt}\cdot\frac{dy}{dt}\cdot\frac{dz}{dt}$

(b)



$$\frac{du}{dt} = \frac{\partial u}{\partial x}\cdot\frac{dx}{dt} + \frac{\partial u}{\partial y}\cdot\frac{dy}{dt} + \frac{\partial u}{\partial z}\cdot\frac{dz}{dt}$$
$$= \nabla u(1,0,1)\cdot(2,1,2)$$
$$= -13/30\ C/s$$

11. A proctologist is walking around the exam room. His position is given by $(x(t), y(t)) = (\cos t, \sin t)$. At position (x, y) his cellphone gets a signal strength of $F(x, y) = e^x y^2$. Using the Chain Rule, find the rate of change of the signal strength with respect to time at $t = \pi/2$.



$z = F(x, y) = e^x y^2$

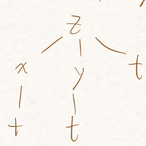
$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\cdot\frac{dx}{dt} + \frac{\partial z}{\partial y}\cdot\frac{dy}{dt}$$
$$= (e^x y^2)\cdot(-\sin t) + (2e^x y)\cdot(\cos t)$$

$t = \frac{\pi}{2}$ $\quad (x(t), y(t)) = (0, 1)$

$$\frac{dz}{dt} = -\sin\left(\frac{\pi}{2}\right) + 2\cos\left(\frac{\pi}{2}\right) = -1$$

12. A particle travels along the path $(x, y) = (t^2 - t, e^{3t})$ in a plane where the temperature at position (x, y) and time t is given by $T(x, y, t) = 2x^2 y \sin t$. Calculate the rate of change of temperature along the particle's path with respect to time at any time t.

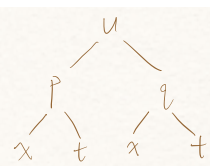
$z = T(x, y, t) = 2x^2 y \sin t$



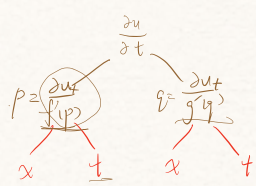
$$\frac{\partial z}{\partial t} = \frac{\partial z}{\partial x}\cdot\frac{dx}{dt} + \frac{\partial z}{\partial y}\cdot\frac{dy}{dt} + \frac{\partial z}{\partial t}$$
$$= (4xy\sin t)(2t-1) + (2x^2\sin t)(3e^{3t}) + 2x^2 y\cdot\cos t$$

$f(x,t)$ $\quad g(x,t)$

13. Show that if f and g are twice differentiable functions, then $u(x, t) = f(x - at) + g(x + at)$ is a solution of the wave equation: $u_{tt} = a^2 u_{xx}$.



$$\frac{\partial u}{\partial t} = \frac{\partial u}{\partial p}\cdot\frac{dp}{dt} + \frac{\partial u}{\partial q}\cdot\frac{dq}{dt}$$
$$= (-a)\,f'(p) + a\,g'(q)$$



$$u_{tt} = (-a)\cdot(-a)\cdot f''(p) + a^2 g''(q)$$
$$u_{xx} = f''(p) + g''(q) \quad \text{QED}$$

14. Let f be a function of two variables and define $g(x,y) = f(\overset{u}{\sin y}, \overset{v}{\cos x})$. Find g_{xx} and g_{yy}. State any assumptions you needed to make.

$f \to u, v$; $u \to y$; $v \to x$

$$\frac{\partial g}{\partial x} = \frac{\partial f}{\partial v}\cdot\frac{\partial v}{\partial x} = \frac{\partial f}{\partial v}(-\sin x)$$

$$\frac{\partial}{\partial x}\left(\frac{\partial g}{\partial x}\right) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial v}(-\sin x)\right)$$

$$= \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial v}\right)(-\sin x) + \left(\frac{\partial f}{\partial v}\right)\underbrace{\frac{\partial}{\partial x}(-\sin x)}_{\cos x}$$

f&g both twice diff'ble

$$\frac{\partial g}{\partial y} = \frac{\partial f}{\partial u}\cdot\frac{\partial u}{\partial y} = \frac{\partial f}{\partial u}(\cos y)$$

$$\frac{\partial}{\partial y}\left(\frac{\partial g}{\partial y}\right) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial u}(\cos y)\right)$$

15. Let $g(u,v) = f(\overset{x}{u^2 - v^2}, \overset{y}{2uv})$ ($z=$). Express $(g_u)^2 + (g_v)^2$ and $g_{uu} + g_{vv}$ in terms of the partial derivatives of f. What hypothesis must f satisfy?

$z \to x, y$; $x \to u, v$; $y \to u, v$

$$g_u = \frac{df}{dx}\frac{dx}{du} + \frac{df}{dy}\cdot\frac{dy}{du} = \frac{df}{dx}(2u) + \frac{df}{dy}2v$$

$$g_v = \frac{df}{dx}\frac{dx}{dv} + \frac{df}{dy}\cdot\frac{dy}{dv} = \frac{df}{dx}(-2v) + \frac{df}{dy}\cdot 2u$$

$$(g_u)^2 + (g_v)^2 = \left(\frac{df}{dx}(2u) + \frac{df}{dy}2v\right)^2 + \left(\frac{df}{dx}(-2v) + \frac{df}{dy}\cdot 2u\right)^2$$

$$= 4(u^2+v^2)(f_x^2 + f_y^2)$$

$$g_{uu} + g_{vv} = \frac{d}{du}\left(\frac{df}{dx}(2u) + \frac{df}{dy}2v\right) + \frac{d}{dv}\left(\frac{df}{dx}(-2v) + \frac{df}{dy}\cdot 2u\right)$$

$$= \frac{d}{du}\frac{df}{dx}\cdot 2u + \cancel{\frac{df}{dx}\cdot 2} + \frac{d}{du}\left(\frac{df}{dy}\right)2v + \frac{df}{dy}\cdot 0$$

$$+ \frac{d}{dv}\left(\frac{df}{dx}\right)(-2v) + \cancel{\frac{df}{dx}(-2)} + \frac{d}{dv}\left(\frac{df}{dy}\right)2u + 0$$

$$= f_u f_x \cdot 2u - f_v f_x 2v + f_u f_y 2v + f_v f_y 2u$$

$$= 4(u^2+v^2)(f_{xx} + f_{yy})$$

咋得ino？

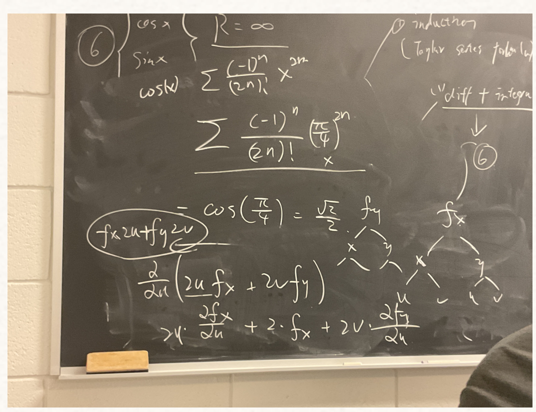
17. A function $g(u)$ with continuous second derivative is given, and f is defined by $f(x,y) = g\left(\frac{x}{y}\right)$, for $y \neq 0$.

Calculate $\frac{\partial^2 f}{\partial x \partial y}$ and $\frac{\partial^2 f}{\partial y \partial x}$ and verify that they are equal.

18. Let $f(x,y) = (xy)^{1/3}$, $p(t) = t$, $q(t) = t^2$, and consider the composite function H defined by $H(t) = f(p(t), q(t))$. Show that the Chain Rule for $H(t)$ is not satisfied at $t = 0$. What conclusion can you draw about f at $(0,0)$?

19. Let $f(x,y,z)$ and $g(x,y,z)$ have continuous partial derivatives. Prove that

$$\nabla(fg) = f\nabla g + g\nabla f$$



20. (a) Let $F(t) = f(a+th, b+tk)$, where the two-variable function f has continuous second partial derivatives, and a, b, h, k are constants. Show that

$$F''(t) = h^2 f_{11} + 2hk f_{12} + k^2 f_{22},$$

where f_{11}, f_{12} and f_{22} are evaluated at $(a + th, b + tk)$.

(b) Can you generalize (a) to give a formula for $F'''(t)$?

21. Functions $f(x, y, z)$ which satisfy Laplace's equation $f_{xx} + f_{yy} + f_{zz} = 0$ are of interest in theoretical physics.

(a) Suppose that the single-variable function g has a continuous second derivative, and $f(x, y, z) = g\left(\frac{1}{r}\right)$, where $r = \sqrt{x^2 + y^2 + z^2} > 0$. Show that

$$f_{xx} + f_{yy} + f_{zz} = \frac{1}{r^4} g''\left(\frac{1}{r}\right), \text{ for } r > 0$$

(b) Give a function f, other than a linear function, which satisfies Laplace's equation.

22. * Let the three-variable function f be differentiable and satisfy $f(t\mathbf{x}) = t^p f(\mathbf{x})$, for all $\mathbf{x} \in \mathbb{R}^3$ and $t \in \mathbb{R}$, where p is constant. Prove that

$$\mathbf{x} \cdot \nabla f(\mathbf{x}) = p f(\mathbf{x}) \quad \text{for all } \mathbf{x} \in \mathbb{R}^3$$

Chapter 7

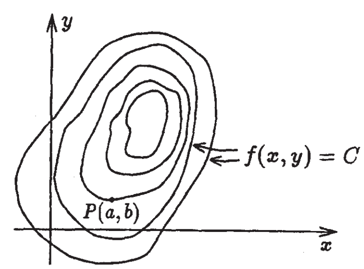
Directional Derivatives and the Gradient Vector

In this chapter, we introduce the concept of the directional derivative of a function. This leads to a geometrical interpretation of the gradient vector.

7.1 Directional Derivatives

Motivation

Let $z = f(x,y)$ represent the height of a mountain. The level curves $f(x,y) = C$ represent the contour lines. Suppose that a skier is at the point $P(a,b)$. In what direction should he move in order to lose height as rapidly as possible?



In order to answer such a question, we have to generalize the idea of the partial derivative. One can think of f_x as the rate of change of f in the x-direction and f_y as the rate of change of f in the y-direction. Our aim is to define a derivative which gives the rate of change of a function f in a direction specified by a unit vector $\hat{u} = (u_1, u_2)$ (i.e. $\|\hat{u}\| = 1$) from a given point (a,b).

If L is the line through (a,b) in the direction $\hat{u}$, then L has vector equation

$$(x,y) = (a,b) + s\hat{u} = (a + su_1, b + su_2), \text{ for } s \in \mathbb{R}$$

At points on the line L, $f(x,y)$ has value $f(a+su_1, b+su_2)$, and this defines a function of one variable s. Thus, the rate of change of f at (a,b) in the direction of $\hat{u}$ is just the derivative of this function with respect to s evaluated at $s = 0$. Hence, we make the following definition.

DEFINITION

Directional Derivative

The **directional derivative** of $f(x, y)$ at a point (a, b) in the direction of a unit vector $\hat{u} = (u_1, u_2)$ is defined by

$$D_{\hat{u}} f(a, b) = \frac{d}{ds} f(a + su_1, b + su_2)\bigg|_{s=0}$$

provided the derivative exists.

REMARK

Letting $g(s) = f(a + su_1, b + su_2)$, we see that we can re-write the definition of $D_{\hat{u}} f(a, b)$ as

$$D_{\hat{u}} f(a, b) = g'(0)$$

By appealing to the definition of the single-variable derivative, we obtain the following alternative expression for the directional derivative:

$$\begin{aligned} D_{\hat{u}} f(a, b) &= \lim_{h \to 0} \frac{g(0 + h) - g(0)}{h} \\ &= \lim_{h \to 0} \frac{f(a + hu_1, b + hu_2) - f(a, b)}{h} \end{aligned}$$

EXAMPLE 1

Find the directional derivative of $f(x, y) = x^2 - y^2$ at the point $(1, 2)$ in the direction of the vector $\vec{u} = \left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)$.

Solution: By definition, we get

$$\begin{aligned} D_{\hat{u}} f(1, 2) &= \frac{d}{ds} f\left(1 + \frac{1}{\sqrt{5}} s, 2 + \frac{2}{\sqrt{5}} s\right)\bigg|_{s=0} \\ &= \frac{d}{ds}\left[\left(1 + \frac{1}{\sqrt{5}} s\right)^2 - \left(2 + \frac{2}{\sqrt{5}} s\right)^2\right]\bigg|_{s=0} \\ &= \left[\frac{2}{\sqrt{5}}\left(1 + \frac{1}{\sqrt{5}} s\right) - \frac{4}{\sqrt{5}}\left(2 + \frac{2}{\sqrt{5}} s\right)\right]\bigg|_{s=0} = -\frac{6}{\sqrt{5}} \end{aligned}$$

The directional derivatives in the directions of $\hat{u} = \hat{i} = (1, 0)$ and $\hat{u} = \hat{j} = (0, 1)$ are familiar objects.

THEOREM 1

(1) $D_{\hat{i}} f = f_x$

(2) $D_{\hat{j}} f = f_y$

Proof: By the remark following the definition of the directional derivative, we find that if $\hat{u} = \hat{i} = (1,0)$ then

$$D_{\hat{i}} f(a,b) = \lim_{h \to 0} \frac{f(a+h,b) - f(a,b)}{h}$$

The right-side is precisely $f_x(a,b)$. The computation with $\hat{u} = \hat{j}$ is similar. □

We now derive a simple formula for calculating the directional derivative of a differentiable function in terms of the partial derivatives.

THEOREM 2

If $f(x,y)$ is differentiable at (a,b) and $\hat{u} = (u_1, u_2)$ is a unit vector, then

$$D_{\hat{u}} f(a,b) = \nabla f(a,b) \cdot \hat{u}$$

Proof: Since f is differentiable at (a,b) we can apply the Chain Rule to get

$$\begin{aligned}
D_{\hat{u}} f(a,b) &= \frac{d}{ds} f(a+su_1, b+su_2)\Big|_{s=0} \\
&= \Big[D_1 f(a+su_1, b+su_2) \frac{d}{ds}(a+su_1) \\
&\quad + D_2 f(a+su_1, b+su_2) \frac{d}{ds}(b+su_2) \Big]\Big|_{s=0} \\
&= \big[D_1 f(a+su_1, b+su_2) u_1 + D_2 f(a+su_1, b+su_2) u_2 \big]\Big|_{s=0} \\
&= D_1 f(a,b) u_1 + D_2 f(a,b) u_2 \\
&= \nabla f(a,b) \cdot (u_1, u_2)
\end{aligned}$$

□

EXAMPLE 2

Find the directional derivative of $f(x,y) = 2x^3 + 4xy^2 + y$ at the point $(-1,1)$ in the direction of the vector $\vec{u} = (1,1)$.

Solution: Observe that the vector is not a unit vector, so we must normalize it. We get

$$\hat{u} = \frac{(1,1)}{\|(1,1)\|} = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right)$$

We have

$$\nabla f(x,y) = (6x^2 + 4y^2, 8xy + 1), \qquad \text{so} \qquad \nabla f(-1,1) = (10, -7)$$

Since f has continuous partial derivatives at $(-1,1)$, it is differentiable at $(-1,1)$. Thus, we can apply Theorem 2 to get

$$D_{\hat{u}} f(-1,1) = (10,-7) \cdot \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right) = \frac{3}{\sqrt{2}}$$

EXAMPLE 3 Find the directional derivative of $f(x, y) = ye^{xy}$ at the point $(2, 1)$ in the direction of the vector $\vec{u} = (-3, 4)$.

Solution: We normalize the vector to get

$$\hat{u} = \frac{(-3, 4)}{\|(-3, 4)\|} = \left(-\frac{3}{5}, \frac{4}{5}\right)$$

We have

$$\nabla f(x, y) = (y^2 e^{xy}, e^{xy} + xye^{xy}), \quad \text{so} \quad \nabla f(2, 1) = (e^2, 3e^2)$$

Since f has continuous partial derivatives at $(2, 1)$, it is differentiable at $(2, 1)$. Thus, we can apply Theorem 2 to get

$$D_{\hat{u}} f(2, 1) = (e^2, 3e^2) \cdot \left(-\frac{3}{5}, \frac{4}{5}\right) = \frac{9e^2}{5}$$

REMARKS

1. Be careful to check the condition of Theorem 2 before applying it. If f is not differentiable at (a, b), then we must apply the definition of the directional derivative.
2. If we choose $\hat{u} = \hat{i} = (1, 0)$ or $\hat{u} = \hat{j} = (0, 1)$, then we find that the directional derivative is equal to the partial derivatives f_x or f_y respectively, just as we had seen in Theorem 1.

The definition of the directional derivative and Theorem 2 can be extended to higher dimensions in the expected way.

EXERCISE 1 Find the directional derivative of f defined by

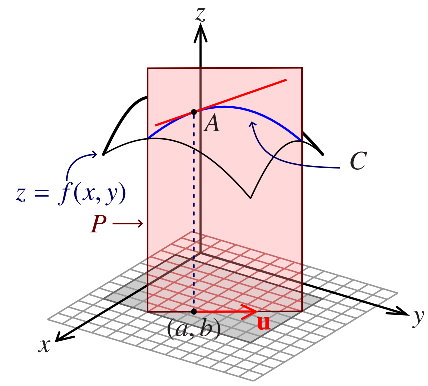
$$f(x, y, z) = e^{xyz}$$

at the point $(1, -1, 2)$ in the direction of the vector $\vec{u} = (1, 2, -2)$.

When the directional derivative is applied, (x, y) usually represents position, and $f(x, y)$ represents some physical quantity, e.g. temperature, or height above sea level. Because the parameter s in the definition represents distance along the line L, the directional derivative represents a rate of change with respect to distance.

For example, if $f(x, y)$ gives the temperature at position (x, y), then $D_{\hat{u}} f(a, b)$ equals the rate of change of temperature, with respect to distance, at position (a, b) in the direction $\hat{u}$, and has dimensions of temperature per unit length.

If $z = f(x, y)$ represents height above sea level, then $D_{\hat{u}} f(a, b)$ equals the rate of change of height z with respect to horizontal distance, at position (a, b) in the direction $\hat{u}$. Geometrically, it equals the slope of the tangent to the cross-section C at the point A. (The vertical plane P cuts the surface $z = f(x, y)$ along the curve C.)



7.2 The Gradient Vector in Two Dimensions

The Greatest Rate of Change

In general, for a function $f(x, y)$, the directional derivative $D_{\hat{u}} f(a, b)$ has infinitely many values corresponding to all possible directions $\hat{u}$ at (a, b). It is natural to ask:

> "In which direction $\hat{u}$ does $D_{\hat{u}} f(a, b)$ assume its largest value?"

This is easily answered using Theorem 7.1.2 and the following property of the dot product:

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

where θ is the angle between $\vec{u}$ and $\vec{v}$.

THEOREM 1 If $f(x, y)$ is differentiable at (a, b) and $\nabla f(a, b) \neq (0, 0)$, then the largest value of $D_{\hat{u}} f(a, b)$ is $\|\nabla f(a, b)\|$, and occurs when $\hat{u}$ is in the direction of $\nabla f(a, b)$.

Proof: Since f is differentiable at (a, b) and $\|\hat{u}\| = 1$ we have

$$\begin{aligned} D_{\hat{u}} f(a, b) &= \nabla f(a, b) \cdot \hat{u} \\ &= \|\nabla f(a, b)\| \, \|\hat{u}\| \cos\theta \\ &= \|\nabla f(a, b)\| \cos\theta \end{aligned}$$

where θ is the angle between $\hat{u}$ and $\nabla f(a, b)$. Thus, $D_{\hat{u}} f(a, b)$ assumes its largest value when $\cos\theta = 1$ i.e. $\theta = 0$. Consequently, the largest value of $D_{\hat{u}} f(a, b)$ is $\|\nabla f(a, b)\|$ and occurs when $\hat{u}$ is in the direction of $\nabla f(a, b)$. $\square$

EXAMPLE 1

Find the largest rate of change of $f(x, y) = \sqrt{x^2 + 2y^2}$ at the point $(1, 2)$, and the direction in which it occurs.

Solution: We have $\nabla f(x, y) = \left(\frac{x}{\sqrt{x^2 + 2y^2}}, \frac{2y}{\sqrt{x^2 + y^2}} \right)$. Thus, by Theorem 1, the largest rate of change of f at $(1, 2)$ is

$$\|\nabla f(1, 2)\| = \left\| \left(\frac{1}{3}, \frac{4}{3} \right) \right\| = \frac{\sqrt{17}}{3}$$

It occurs in the direction

$$\vec{u} = \nabla f(1, 2) = \left(\frac{1}{3}, \frac{4}{3} \right)$$

EXAMPLE 2

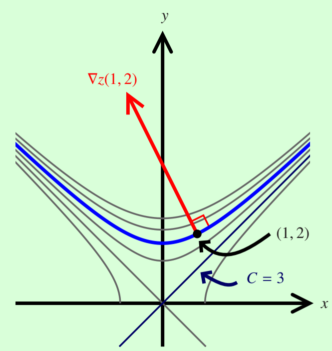
Let $z = f(x, y) = 3 - x^2 + y^2$ represent the height above sea level. A hiker is at position $(1, 2, 6)$. In what direction should he start to move in order to follow a path of steepest ascent? What would be the slope of his path (i.e. rate of change of height with respect to horizontal distance)?

Solution: The gradient of f is

$$\nabla f(x, y) = (-2x, 2y)$$

and at the given point

$$\nabla f(1, 2) = (-2, 4)$$



By Theorem 1, the hiker should move in the direction $\vec{u} = (-2, 4)$ in order to follow a path of steepest ascent (i.e. largest rate of change of f). The slope of his path would be

$$\|\nabla f(1, 2)\| = \sqrt{(-2)^2 + (4)^2} = 2\sqrt{5}$$

EXERCISE 1

Find the largest rate of change of $f(x, y) = \ln(x + y^2)$ at the point $(0, 1)$, and the direction in which it occurs.

EXERCISE 2

Give a non-constant function $f(x, y)$ and a point (a, b) such that the directional derivative at (a, b) is independent of the direction. What can you say about the tangent plane of the surface $z = f(x, y)$ at the point (a, b)?

Theorem 1 also applies in any dimension. That is, if $f(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^n$, is differentiable at $\mathbf{a}$ and $\hat{u} \in \mathbb{R}^n$ is a unit vector, then the largest value of $D_{\hat{u}} f(\mathbf{a})$ is $\|\nabla f(\mathbf{a})\|$, and it occurs when $\hat{u}$ is in the direction of $\nabla f(\mathbf{a})$.

EXAMPLE 3 Let $f(x, y, z) = z^3 e^{x^2+y^2-2x}$. Determine the greatest rate of change of f at $(1, 1, 1)$ and the direction in which it occurs.

Solution: We have

$$\nabla f = \left((2x-2)z^3 e^{x^2+y^2-2x}, 2yz^3 e^{x^2+y^2-2x}, 3z^2 e^{x^2+y^2-2x}\right)$$

Thus, the greatest rate of change of f at $(1, 1, 1)$ is

$$\|\nabla f(1, 1, 1)\| = \|(0, 2, 3)\| = \sqrt{0+4+9} = \sqrt{13}$$

and occurs in the direction of

$$\vec{u} = \nabla f(1, 1, 1) = (0, 2, 3)$$

The Gradient and the Level Curves of f

People who have experience reading contour maps know that the direction of steepest ascent is orthogonal to the contour lines. In mathematical terms, this means that the direction of greatest rate of change of f, which we have shown is the direction of the gradient of f, is orthogonal to the level curves of f. We can see this result geometrically in Example 2. Indeed, as we move along a level curve, the value of f doesn't change, so its rate of change in this direction is zero. Therefore, if $\hat{u}$ is a unit vector pointing in the direction of the level curve at (a, b), we have $D_{\hat{u}} f(a, b) = 0$ and consequently $\nabla f(a, b) \cdot \hat{u} = 0$. That is, ∇f is orthogonal to the direction of the level curve.

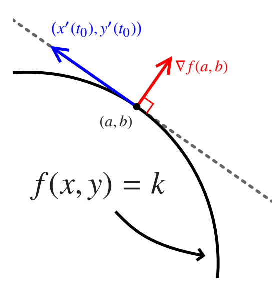
一阶导连续

We now derive this result analytically.

THEOREM 2 If $f(x, y) \in C^1$ in a neighborhood of (a, b) and $\nabla f(a, b) \neq (0, 0)$, then $\nabla f(a, b)$ is orthogonal to the level curve $f(x, y) = k$ through (a, b).

Proof: Since $\nabla f(a, b) \neq (0, 0)$, by the Implicit Function Theorem (see Appendix A) the level curve $f(x, y) = k$ can be described by parametric equations $x = x(t)$, $y = y(t)$ for $t \in I$ where $x(t)$ and $y(t)$ differentiable. Hence, the level curve may be written as $f(x(t), y(t)) = k$, $t \in I$. Suppose

$$a = x(t_0), \quad b = y(t_0) \quad \text{for some } t_0 \in I$$



Since f is differentiable, we can take the derivative of this equation with respect to t using the Chain Rule to get

$$f_x(x(t), y(t))x'(t) + f_y(x(t), y(t))y'(t) = 0$$

On setting $t = t_0$ we get

$$\nabla f(a, b) \cdot (x'(t_0), y'(t_0)) = 0$$

Thus, $\nabla f(a, b)$ is orthogonal to $(x'(t_0), y'(t_0))$ which is tangent to the level curve. □

EXERCISE 3 Prove that the level curves of the functions f and g defined by

$$f(x, y) = \frac{y}{x^2}, \quad x \neq 0, \qquad g(x, y) = x^2 + 2y^2$$

intersect orthogonally. Illustrate graphically.

The Gradient Vector Field

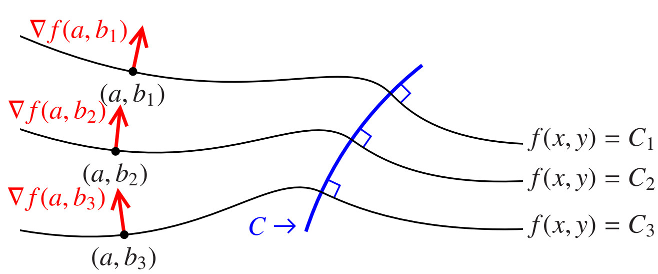
Given a function $f(x, y)$ that is differentiable at (x, y), the gradient of f at (x, y) is defined by

$$\nabla f(x, y) = (f_x(x, y), f_y(x, y))$$

The gradient of f associates a vector with each point of the domain of f, and is referred to as a **vector field**. It is represented graphically by drawing $\nabla f(a, b)$ as a vector emanating from the corresponding point (a, b).

Theorems 1 and 2 show that the gradient vector field has important geometric properties:

1) It gives the direction in which the function has its largest rate of change.
2) It gives the direction that is orthogonal to the level curves of the function.



If the level curves are contour lines, then a curve such as C, which intersects the level curves orthogonally, would define a curve of steepest ascent on the surface.

REMARK

Vector fields and gradient vector fields will be studied in detail in Calculus 4 (AMath 231).

7.3 The Gradient Vector in Three Dimensions

One cannot visualize the graph $w = f(x, y, z)$ of a function $f(x, y, z)$, because four dimensions are required. One can gain insight into such a function, however, by considering the level surfaces in $\mathbb{R}^3$ defined by

$$f(x, y, z) = k, \qquad \text{where } k \in R(f)$$

EXAMPLE 1 The level surfaces of the function f defined by

$$f(x, y, z) = x + 2y + 3z$$

are the parallel planes

$$x + 2y + 3z = k$$

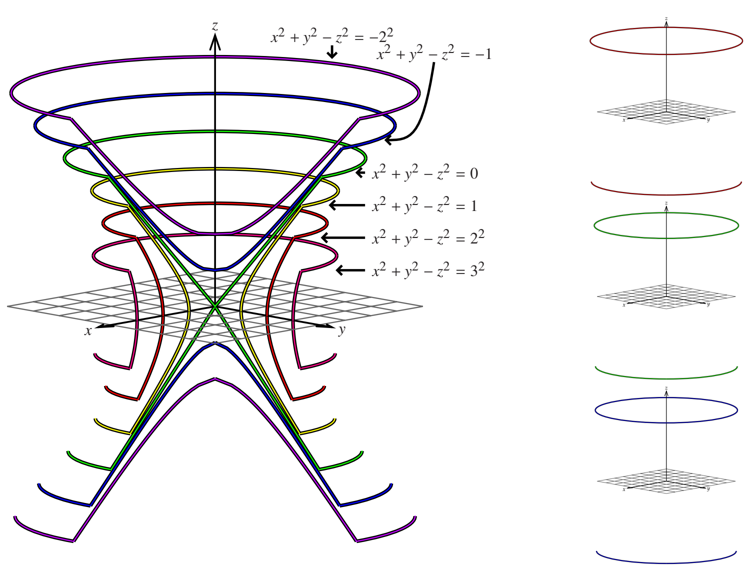
EXAMPLE 2 The level surfaces of the function

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

given by

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

are hyperboloids with two sheets if $k < 0$, hyperboloids with one sheet if $k > 0$, and a cone if $k = 0$.



We now discuss the interpretation of the gradient $\nabla f(a,b,c)$, for $f(x,y,z)$. As noted in Section 7.2, Theorem 7.2.1 applies in this case. That is, $\nabla f(a,b,c)$ gives the direction of the largest rate of change of f. We now generalize Theorem 7.2.2 to the case $f(x,y,z)$. As one might guess, we have:

THEOREM 1 If $f(x,y,z) \in C^1$ in a neighborhood of (a,b,c) and $\nabla f(a,b,c) \neq (0,0,0)$, then $\nabla f(a,b,c)$ is orthogonal to the level surface $f(x,y,z) = k$ through (a,b,c).

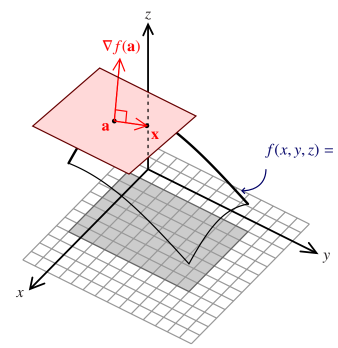
The details are similar to the proof of Theorem 7.2.2.

Observe that Theorem 1 gives a quick way to find the equation of the tangent plane of a surface in $\mathbb{R}^3$ given by

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

If $\mathbf{x} \in \mathbb{R}^3$ is an arbitrary point in the tangent plane to the surface at the point $\mathbf{a} \in \mathbb{R}^3$, then the vector $\mathbf{x} - \mathbf{a}$ lies in the tangent plane, and by Theorem 1, is orthogonal to $\nabla f(\mathbf{a})$, leading to

$$\nabla f(\mathbf{a}) \cdot (\mathbf{x} - \mathbf{a}) = 0$$



Since this equation is satisfied for all $\mathbf{x}$ in the tangent plane, it is the equation of the tangent plane. In component form, we have

$$f_x(a,b,c)(x-a) + f_y(a,b,c)(y-b) + f_z(a,b,c)(z-c) = 0$$

EXAMPLE 3 Find the equation of the tangent plane to the surface $z^3 e^{x^2+y^2-2x} = 1$ at the point $(1,1,1)$.

Solution: From our work above, the equation of the tangent plane is

$$\nabla f(1,1,1) \cdot (x-1, y-1, z-1) = 0$$

Using our work in Example 7.2.3, we get

$$\begin{aligned} (0,2,3) \cdot (x-1, y-1, z-1) &= 0 \\ 2(y-1) + 3(z-1) &= 0 \end{aligned}$$

EXERCISE 1 Find the equation of the tangent plane to the ellipsoid $x^2 + 2y^2 + 3z^2 = 12$ at the point $(1, 1, \sqrt{3})$.

Here is an important special case of the above set-up. Consider the level surface $f(x,y,z) = k$ for a function $f(x,y,z)$ of the form $f(x,y,z) = z - g(x,y)$ where $g(x,y)$

is a function of two variables. Suppose further that $k = 0$. That is, our level surface is of the form

$$z - g(x, y) = 0$$

or more simply

$$z = g(x, y).$$

Let's derive the equation of the tangent plane at (a, b, c) using the above formula. Note first that if (a, b, c) is on the level surface, i.e. if $f(a, b, c) = 0$, then we must have $c = g(a, b)$.

Now, since $f(x, y, z) = z - g(x, y)$, we find that $f_x = -g_x$, $f_y = -g_y$, and $f_z = 1$.

Consequently, the equation for the tangent plane to the surface $z = g(x, y)$ at the point $(a, b, f(a, b))$ is given by

$$-g_x(a, b)(x - a) - g_y(a, b)(y - b) + z - g(a, b) = 0$$

which we can re-write as

$$z = g(a, b) + g_x(a, b)(x - a) + g_y(a, b)(y - b)$$

This is precisely the same expression we obtained in Chapter 5! The utility of this new point of view is that it may be applied to obtain equations for tangent planes of surfaces that are not defined by equations of the form $z = g(x, y)$, where z is given explicitly in terms of the variables x and y.

EXERCISE 2 Find the equation of the tangent plane to the surface

$$z = \frac{xy}{3x - 2y} \quad \text{at } (1, 2, -2)$$

Hint: Rewrite the equation as $z(3x - 2y) - xy = 0$ and use the above approach.

EXERCISE 3 Find the equation of the tangent plane to the surface defined by $x^5 y + y^5 z + z^5 x = -1$ at the point $(1, -1, 1)$.

Chapter 7 Problem Set

1. (a) Calculate the directional derivative of f at the point (a, b) in the direction defined by $\vec{v}$:
   (i) $f(x, y) = e^x \cos y$, $(a, b) = \left(0, \frac{\pi}{4}\right)$, and $\vec{v} = (1, 3)$.
   (ii) $f(x, y, z) = \sin(xyz)$, $(a, b, c) = \left(1, 1, \frac{\pi}{4}\right)$, and $\vec{v} = (1, -\sqrt{2}, 1)$.

   (b) In each case find the direction at (a, b) in which the rate of change of f is greatest, and find this maximum rate of change.
2. The temperature of a metal sheet as a function of position (x, y) is given by $T(x, y) = 100 + 10e^{-x} \sin y$. Find the rate of change of temperature at the point $(0, \frac{\pi}{4})$ in the direction of the vector $(1, 1)$. Find the direction at $(0, \frac{\pi}{4})$ in which the rate of change is greatest, and find this rate of change.
3. Calculate the directional derivative of $g(x, y, z) = \ln(x + e^{yz})$ at $(0, 1, 0)$ in the direction from the point $(0, 1, 0)$

to the point $(2, 3, -1)$.

4. Let $f(x, y) = \ln(x + 2y)$. Find the directional derivative of f at (1,0) in the direction of the line $y = 2x - 2$.

5. Let $f(x, y) = \ln(x^2 + y^2)$.
   (a) Find the directional derivative of f at $(-1, 2)$ in the direction of the vector $\vec{v} = (3, -4)$.
   (b) Find the direction in which f is increasing the fastest at $(1, 1)$. What is the magnitude of this rate of change?
   (c) Find the equation of the tangent line at $(1, 1)$ to the level curve $f(x, y) = \ln 2$.

6. Let $f(x, y) = 2xy - y^2$. Use the gradient vector to find the equation of the tangent line of the curve $f(x, y) = 3$ at the point $(2, 1)$. Sketch the curve and the tangent line.

7. Let $f(x, y, z)$ be a differentiable function such that $\nabla f(a, b, c) \neq (0, 0, 0)$. Consider the surface $f(x, y, z) = k$ and assume that $f(a, b, c) = k$. Write down the equation of the tangent plane to the surface at (a, b, c), in terms of the gradient vector.

8. Let $f(x, y, z) = x^2 + 2y^2 - 3z^2$. Use the gradient vector to find the equation of the tangent plane to the surface $f(x, y, z) = 3$ at the point $(2, 1, 1)$.

9. Let $x^2 - y^2 + 3z^2 = 0$ implicitly define a surface. Find the equation of the tangent plane to the surface at the point $(1, 2, 1)$.

10. Use the gradient vector to verify that the two families of curves intersect each other orthogonally. Illustrate graphically.
    (a) $xy = c$ and $y^2 - x^2 = k$
    (b) $(x - c)^2 + y^2 = c^2$ and $x^2 + (y - k)^2 = k^2$.

11. A sphere centered at $(2, 1, -1)$ passes through the point $P = (1, -1, 1)$. Find the equation of the tangent plane to the sphere at P. Sketch the sphere and plane.

12. (a) Find the directional derivative of $w = x^2 + y^2$ in the direction of the tangent vector to the spiral $(x, y) = (e^t \cos t, 2e^t \sin t)$, at the point defined by $t = 0$.
    (b) Find $\dfrac{dw}{dt}$ along the spiral, at the same point.
    (c) How are these rates of change related?

13. At a point $(a, b) \in \mathbb{R}^2$, the directional derivative of a differentiable function $f(x, y)$ in the directions $(1, 1)$ and $(1, -1)$ equals 3 and 2 respectively. Find the largest rate of change of $f(x, y)$ at (a, b), and the direction in which it occurs.

14. In what directions at the point (2,1) does the directional derivative of the function $f(x, y) = xy$ equal 0? Equal $\sqrt{\frac{5}{2}}$? Express your answer by giving the angle between the required directions and the gradient of f at (2,1). Give a diagram, showing some typical level curves of f near (2,1), and the required directions.

→后面没写

15. The temperature in a region of space is given by $T(x, y, z) = e^{-2x}(1 + 2y)(\frac{1}{1+3z})$. A fly moves along the path $(x, y, z) = (2t, \sin t, e^t - 1)$.
    (a) Find $\frac{dT}{dt}$ at $t = 0$.
    (b) Observe that the direction of the fly's path at $t = 0$ is $(2, 1, 1)$. Find the directional derivative of T in the direction of the fly's path at $t = 0$.
    (c) Explain the physical difference between a) and b).

16. Suppose that another bug, starting at the origin, is about to fly in the temperature field from the previous question. This new bug always flies at a speed of 2 m/s (suppose that the spatial units are metres and the temporal unit is seconds). In what direction(s) could the bug fly initially to experience a rate of change of temperature of 8 °C per second?

17. A space-ship cruising on the sunny side of the planet Mercury starts to overheat. The space-ship is at location (1,1,1) and the temperature of the ship's hull when at location (x, y, z) will be
$$T(x, y, z) = 200 + e^{-x^2-2y^2-3z^2}$$
where x, y, z are in metres.
    (a) In what direction should the ship proceed in order to decrease temperature most rapidly?
    (b) If the ship travels at e^8 m/sec, how fast will the temperature decrease (in degrees/sec) if it proceeds in that direction?
    (c) The metal of the hull will crack if cooled at a rate greater than $\sqrt{14}e^2$ degrees/sec. Describe the set of possible directions in which the ship may proceed to bring the temperature down at that rate. Give a sketch.

18. Let $g(x, y, z) = xe^y + yz^2$. Use the gradient vector to find the equation of the tangent plane to the surface $g(x, y, z) = 2$ at the point $(2, 0, 1)$.

19. Find all points on the paraboloid $z = x^2 + y^2 - 1$ at which the normal line to the surface coincides with the line joining the origin to the point. Illustrate your results with a sketch.

20. A cone, with vertex $(0, 0, -2)$ and axis the z-axis, intersects the plane $z = 3$ in a circle of radius $\sqrt{5}$.
    (a) Show that the tangent plane to the cone at the point $(1, -2, 3)$ cuts the x-axis at the point (2,0,0). Give a sketch.
    (b) Write down a vector equation for the normal line to the cone at (1,-2,3). Hence show that this line intersects the xy-plane at the point (4,-8,0).

21. (*Chemotaxis*) Chemotaxis is the chemically directed movement of organisms up a concentration gradient. The slime mold *Dictyostelium discoideum* exhibits this phenomenon. In this case, single-celled amoeba of this species move up the concentration gradient of a chemical called cyclic AMP. Suppose the concentration of cyclic AMP at the point (x, y) is given by

$$f(x, y) = \frac{4}{xy+1}$$

   (a) If you place an amoeba at the point (3,1) in the xy-plane, determine in which direction the amoeba will move if its movement is directed by chemotaxis.

   (b) It can be shown in general that a particle moving in the manner described above has path $y = y(x)$ satisfying the differential equation (DE)

$$\frac{dy}{dx} = f_y/f_x$$

   Find the path of the amoeba in a), which has initial condition $y(3) = 1$.

22. (a) Consider the sphere of radius 4 centered at the origin, and the sphere of radius 3 centered at the point (0,5,0). Prove that the normal directions to these spheres at their points of intersection are orthogonal. Give a sketch.

   (b) Generalize this result.

23. An engineer wishes to build a railroad up a mountain that has the shape of an elliptic paraboloid $z = c - ax^2 - by^2$, where a, b, c are positive constants. At the point (1,1), in what directions may the track be laid so that it will be climbing with a slope of 0.03 (i.e. a vertical rise of 0.03m for each horizontal metre)? Make a sketch showing a few level curves, the gradient ∇z at (1,1), and the two possible directions for the track. Work out the details using $a = \sqrt{3}b$, $b = 0.015$.

24. Compute the directional derivative $D_{\vec{u}} f(0, 0)$ in the direction $\vec{u} = (u_1, u_2)$ of

$$f(x, y) = \begin{cases} \dfrac{x|y|}{\sqrt{x^2 + y^2}} & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{if } (x, y) = (0, 0) \end{cases}$$

25. Let $f(x, y)$ be differentiable for all $(x, y) \in \mathbb{R}^2$. If $f(x, y) = f(y, x)$ for all (x, y), prove that the directional derivative of f at $(0, 0)$ in the direction $\left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$ is 0.

1. (a) Calculate the directional derivative of f at the point (a,b) in the direction defined by $\vec{v}$:

(i) $f(x,y) = e^x \cos y$, $(a,b) = \left(0, \frac{\pi}{4}\right)$, and $\vec{v} = (1,3)$.

(ii) $f(x,y,z) = \sin(xyz)$, $(a,b,c) = \left(1, 1, \frac{\pi}{4}\right)$, and $\vec{v} = (1, -\sqrt{2}, 1)$.

(b) In each case find the direction at (a,b) in which the rate of change of f is greatest, and find this

(i) $\|(1,3)\| = \sqrt{10}$

$\nabla f(x,y) = (e^x \cos y, -e^x \sin y)$

$\nabla f(0, \frac{\pi}{4}) = (\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2})$

$\because$ f cont at $(0, \frac{\pi}{4})$

$\therefore$ $D_{\vec{v}} f(a,b) = \frac{1}{\sqrt{10}} (\frac{\sqrt{2}}{2}, -\frac{\sqrt{2}}{2})(1,3)$

$= -\frac{1}{\sqrt{5}}$

max $\|\nabla f(a,b)\| = \sqrt{(\frac{\sqrt{2}}{2})^2 + (\frac{\sqrt{2}}{2})^2} = 1$

(ii) $\|(1, -\sqrt{2}, 1)\| = \sqrt{1+2+1} = 2$

$\nabla f(x,y,z) = (yz\cos(xyz), xz\cos(xyz), xy\cos(xyz))$

$\nabla f(1,1,\frac{\pi}{4}) = (\frac{\sqrt{2}}{8}\pi, \frac{\sqrt{2}}{8}\pi, \frac{\sqrt{2}}{2})$

$\because f$ cont at $(1, 1, \frac{\pi}{4})$

$\therefore$ by DD theorem

$D_{\vec{v}} f(a,b,c) = \frac{1}{2}(1, 1, \frac{\pi}{4})(1, -\sqrt{2}, 1)$

$= \frac{\sqrt{2}}{16}\pi - \frac{1}{8}\pi + \frac{\sqrt{2}}{4}$

max $\|\nabla f(a,b,c)\| = \sqrt{(\frac{\sqrt{2}}{8}\pi)^2 + (\frac{\pi}{4})^2 + \frac{1}{2}}$

2. The temperature of a metal sheet as a function of position (x,y) is given by $T(x,y) = 100 + 10e^{-x}\sin y$. ① Find the rate of change of temperature at the point $(0, \frac{\pi}{4})$ in the direction of the vector $(1,1)$. ② Find the direction at $(0, \frac{\pi}{4})$ in which the rate of change is greatest, and find this rate of change.

① $\|(1,1)\| = \sqrt{2}$

$\nabla T(x,y) = (-10e^{-x}\sin y, 10e^{-x}\cos y)$

$\nabla T(0, \frac{\pi}{4}) = (-5\sqrt{2}, 5\sqrt{2})$

$\because T$ cont at $(0, \frac{\pi}{4})$

$\therefore$ By DD theorem.

$D_{\vec{v}} T(0, \frac{\pi}{4}) = \frac{1}{\sqrt{2}}(1,1)(-5\sqrt{2}, 5\sqrt{2}) = 0$

② By greatest rate of Change Theorem.

max is $\|\nabla T(0, \frac{\pi}{4})\| = 10$

3. Calculate the directional derivative of $g(x,y,z) = \ln(x + e^{yz})$ at $(0,1,0)$ in the direction from the point $(0,1,0)$ to the point $(2,3,-1)$.

direction vector : $\vec{v} = (2, 2, -1)$

$\|(2,2,-1)\| = \sqrt{4+4+1} = 3$

$\nabla f(x,y,z) = (\frac{1}{x+e^{yz}}, \frac{ze^{yz}}{x+e^{yz}}, \frac{ye^{yz}}{x+e^{yz}})$

$\nabla f(0,1,0) = (1, 0, 1)$

$\because f$ cont at $(0,1,0)$

$\therefore$ By DD thoren

$D_{\vec{v}} f(2,3,-1) = \frac{1}{3}(1,0,1)(2,3,-1)$

$= \frac{1}{6}$

4. Let $f(x,y)=\ln(x+2y)$. Find the directional derivative of f at (1,0) in the direction of the line $y=2x-2$.

direction vector: $\vec{u}=(2,2)-(1,0)=(1,2)$

$\|\hat{u}\|=\pm\sqrt{5}$ $\quad \vec{u}=(1,0)-(2,2)=-(1,2)$

$\nabla f(x,y)=\left(\frac{1}{x+2y},\frac{2}{x+2y}\right)$

$\nabla f(1,0)=(1,2)$

$D_{\hat{u}}f(1,0)=\pm\frac{1}{\sqrt{5}}(1,2)(1,2)$

$=\pm\sqrt{5}$

5. Let $f(x,y)=\ln(x^2+y^2)$.

(a) Find the directional derivative of f at $(-1,2)$ in the direction of the vector $\vec{v}=(3,-4)$.

(b) Find the direction in which f is increasing the fastest at $(1,1)$. What is the magnitude of this rate of change?

(c) Find the equation of the tangent line at $(1,1)$ to the level curve $f(x,y)=\ln 2$.

(a) $\|\vec{v}\|=\sqrt{9+16}=5$

$\nabla f(x,y)=\left(\frac{2x}{x^2+y^2},\frac{2y}{x^2+y^2}\right)$

$\nabla f(-1,2)=\left(\frac{-2}{5},\frac{4}{5}\right)$

$D_{\hat{v}}f(-1,2)=\frac{1}{5}\left(-\frac{2}{5},\frac{4}{5}\right)(3,-4)$

$=\frac{1}{5}\left(-\frac{6}{5}-\frac{16}{5}\right)$

$=\frac{-22}{25}$

(b) direction $(1,1)$

$\|\nabla f(1,1)\|=\left\|\left(\frac{2}{2},\frac{2}{2}\right)\right\|=\sqrt{2}$

(c) ∵ $f(1,1)=\ln 2$

∴ By Orthogonality theo, $\nabla f(1,1)\perp f(x,y)=\ln 2$

$\nabla f(1,1)(x-1,y-1)=0$

$(1,1)(x-1,y-1)=0$

$(x-1)+(y-1)=0$

6. Let $f(x,y)=2xy-y^2$. Use the gradient vector to find the equation of the tangent line of the curve $f(x,y)=3$ at the point $(2,1)$. Sketch the curve and the tangent line.

∵ $f(2,1)=3$

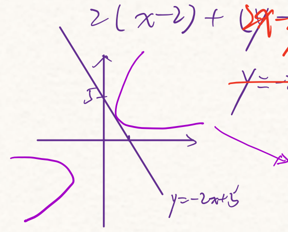
∴ By orthogonality theo. $\nabla f(2,1)\perp f(x,y)=3$.

$\nabla f(2,1)(2y,2x-2y)=0$

$2(x-2)+\cancel{(y-1)}=0$

~~$y=-2x+5$~~

小错



curve: $2xy-y^2=3$

7. Let $f(x,y,z)$ be a differentiable function such that $\nabla f(a,b,c)\neq(0,0,0)$. Consider the surface $f(x,y,z)=k$ and assume that $f(a,b,c)=k$. Write down the equation of the tangent plane to the surface at (a,b,c), in terms of the gradient vector.

$\nabla f(a,b,c)\cdot(x-a,y-b,z-c)=0$

8. Let $f(x,y,z)=x^2+2y^2-3z^2$. Use the gradient vector to find the equation of the tangent plane to the surface $f(x,y,z)=3$ at the point $(2,1,1)$.

∵ $f(2,1,1)=4+2-3=3$

∴ By orthonality theorem.

$\nabla f(x,y,z)(x-2,y-1,z-1)=0$

$f_x(2,1,1)=2x=4$

$f_y(2,1,1)=4y=4$

$f_z(2,1,1)=-6z=-6$

$4(x-2)+4(y-1)-6(z-1)=0$

9. Let $x^2 - y^2 + 3z^2 = 0$ implicitly define a surface. Find the equation of the tangent plane to the surface at the point $(1, 2, 1)$.

Let $f(x,y,z) = x^2 - y^2 + 3z^2$

$\because f(1,2,1) = 0$

$\therefore$ by orthogonality theo

$\nabla f(x,y,z)(x-1, y-2, z-1) = 0$

$f_x(1,2,1) = 2x = 2$

$f_y(1,2,1) = -2y = -4$

$f_z(1,2,1) = 6z = 6$

$2(x-1) - 4(y-2) + 6(z-1) = 0$

$2x - 2 - 4y + 8 + 6z - 6 = 0$

$6z = -2x + 4y$

$z = -\frac{1}{3}x + \frac{2}{3}y$

10. Use the gradient vector to verify that the two families of curves intersect each other orthogonally. Illustrate graphically.

(a) $xy = c$ ① and $y^2 - x^2 = k$ ②

(b) $(x - c)^2 + y^2 = c^2$ and $x^2 + (y - k)^2 = k^2$.

(a) → 有交点

$y = \frac{c}{x}$ 代入②

$\frac{c^2}{x^2} - x^2 = k$.

11. A sphere centered at $(2, 1, -1)$ passes through the point $P = (1, -1, 1)$. Find the equation of the tangent plane to the sphere at P. Sketch the sphere and plane.

General equation of sphere centred at (a,b,c)

$(x-a)^2 + (y-b)^2 + (z-c)^2 = r^2$

→ $(x-2)^2 + (y-1)^2 + (z+1)^2 = r^2$

$r^2 = (1-2)^2 + (-1-1)^2 + (1+1)^2 = 9$ $\quad r = 3$

→ let $f(x,y,z) = (x-2)^2 + (y-1)^2 + (z+1)^2$

$\nabla f(x,y,z) = (2x-4, 2y-2, 2z+2)$

$\nabla f(1,-1,1) = (2, -4, 4)$

$\nabla f(1,-1,1) \cdot (x-1, y+1, z-1) = 0$

$-2(x-1) - 4(y+1) + 4(z-1) = 0$

12. (a) Find the directional derivative of $w = x^2 + y^2$ in the direction of the tangent vector to the spiral $(x, y) = (e^t \cos t, 2e^t \sin t)$, at the point defined by $t = 0$.

(b) Find $\frac{dw}{dt}$ along the spiral, at the same point.

(c) How are these rates of change related?

(a)

13. At a point $(a, b) \in \mathbb{R}^2$, the directional derivative of a differentiable function $f(x, y)$ in the directions $(1, 1)$ and $(1, -1)$ equals 3 and 2 respectively. Find the largest rate of change of $f(x, y)$ at (a, b), and the direction in which it occurs.

$u = (1,1) \quad \|u\| = \sqrt{2}$

$v = (1,-1) \quad \|v\| = \sqrt{2}$

$D_u f(a,b) = \frac{1}{\sqrt{2}} \nabla f(a,b)(1,1) = 3$

$D_v f(a,b) = \frac{1}{\sqrt{2}} \nabla f(a,b)(1,-1) = 2.$

$\nabla f(a,b)(1,1) = 3\sqrt{2}$

$\nabla f(a,b)(1,-1) = 2\sqrt{2}$

$\begin{cases} f_x - f_y = 3\sqrt{2} \\ f_x - f_y = 2\sqrt{2} \end{cases}$

$f_x = \frac{5\sqrt{2}}{2} \quad f_y = -\frac{\sqrt{2}}{2}$

$\|\nabla f(a,b)\| = \sqrt{\left(\frac{5\sqrt{2}}{2}\right)^2 + \left(\frac{\sqrt{2}}{2}\right)^2}$

direction $(f_x, f_y) = \left(\frac{5\sqrt{2}}{2}, -\frac{\sqrt{2}}{2}\right)$

14. In what directions at the point (2,1) does the directional derivative of the function $f(x, y) = xy$ equal 0? Equal $\sqrt{\frac{5}{2}}$? Express your answer by giving the angle between the required directions and the gradient of f at (2,1). Give a diagram, showing some typical level curves of f near (2,1), and the required directions.

Chapter 8

Taylor Polynomials and Taylor's Theorem

For a function of one variable f, the second derivative f'' plays an important role in approximating $f(x)$. Geometrically, f'' determines whether the graph of f is concave up or concave down. Thus, if the graph of f is concave up near x ($f''(x) > 0$), then the linear approximation formula gives a value for $f(x)$ which is too small. The second derivative can in fact be used to estimate the error through Taylor's formula. In addition, f'' can be used to increase the accuracy of the linear approximation by defining a quadratic approximation, the second degree Taylor polynomial.

In this chapter, we extend these ideas to functions of two variables.

8.1 The Taylor Polynomial of Degree 2

Review of the 1-D case

For a function of one variable, $f(x)$, the Taylor polynomial of degree 2 at a is denoted by $P_{2,a}(x)$, and is defined by

$$P_{2,a}(x) = f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2$$

Observe that $P_{2,a}(x)$ is the sum of the linear approximation $L_a(x)$ and a term which is of second degree in $(x-a)$. The coefficient of this term is determined by requiring that the second derivative of $P_{2,a}(x)$ equals the second derivative of f at a:

$$P''_{2,a}(a) = f''(a)$$

You should verify this by differentiating $P_{2,a}(x)$.

The 2-D case

Suppose that $f(x, y)$ has continuous second partials at (a, b). The Taylor polynomial of f of degree 2 at (a, b) is denoted $P_{2,(a,b)}(x, y)$ and is obtained by adding appropriate 2nd degree terms in $(x-a)$ and $(y-b)$ to the linear approximation $L_{(a,b)}(x, y)$. Consider

$$P_{2,(a,b)}(x, y) = L_{(a,b)}(x, y) + A(x-a)^2 + B(x-a)(y-b) + C(y-b)^2 \tag{8.1}$$

where A, B, C are constants. Using (8.1) we can find that

$$\frac{\partial^2 P_{2,(a,b)}}{\partial x^2} = 2A$$

since $L_{(a,b)}(x, y)$ does not contribute to the second derivatives as it is of first degree in x and y.

Similarly, finding the other second partial derivatives of $P_{2,(a,b)}(x, y)$ gives

$$\frac{\partial^2 P_{2,(a,b)}}{\partial x \partial y} = B$$
$$\frac{\partial^2 P_{2,(a,b)}}{\partial y^2} = 2C$$

Requiring that the second partial derivatives of $P_{2,(a,b)}$ equal the second partial derivatives of f at (a, b) leads to

$$2A = \frac{\partial^2 f}{\partial x^2}(a, b), \quad B = \frac{\partial^2 f}{\partial x \partial y}(a, b), \quad 2C = \frac{\partial^2 f}{\partial y^2}(a, b)$$

We then substitute these into equation (8.1) and write out the expression for $L_{(a,b)}(x, y)$, to obtain the required formula.

DEFINITION

2nd degree Taylor polynomial

The **second degree Taylor polynomial** $P_{2,(a,b)}$ of $f(x, y)$ at (a, b) is given by

$$\begin{aligned} P_{2,(a,b)}(x, y) &= f(a, b) + f_x(a, b)(x-a) + f_y(a, b)(y-b) \\ &\quad + \frac{1}{2}\Big[f_{xx}(a, b)(x-a)^2 + 2f_{xy}(a, b)(x-a)(y-b) + f_{yy}(a, b)(y-b)^2\Big] \end{aligned}$$

In general, it approximates $f(x, y)$ for (x, y) sufficiently close to (a, b):

$$f(x, y) \approx P_{2,(a,b)}(x, y)$$

with better accuracy than the linear approximation.

EXAMPLE 1 Use the Taylor polynomial of degree 2 to calculate $\sqrt{(0.95)^3 + (1.98)^3}$ approximately. [This is a continuation of Example 4.4.1.]

Solution: Let $f(x, y) = \sqrt{x^3 + y^3}$ and $(a, b) = (1, 2)$. By differentiating, one obtains

$$\nabla f(1, 2) = \left(\frac{1}{2}, 2\right), \qquad Hf(1, 2) = \begin{bmatrix} \frac{11}{12} & -\frac{1}{3} \\ -\frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

Thus,

$$P_{2,(1,2)}(x, y) = 3 + \frac{1}{2}(x-1) + 2(y-2) + \frac{1}{2}\left[\frac{11}{12}(x-1)^2 - \frac{2}{3}(x-1)(y-2) + \frac{2}{3}(y-2)^2\right]$$

This polynomial approximates $\sqrt{x^3 + y^3}$ near the point $(1, 2)$:

$$\begin{aligned} \sqrt{(0.95)^3 + (1.98)^3} &\approx P_{2,(1,2)}(0.95, 1.98) \\ &= 3 + (-0.065) + \left(\frac{0.0227}{12}\right) \\ &= 2.935946 \end{aligned}$$

The calculator value is 2.935944. Hence, the error is 0.000002 compared with 0.000943 for the linear approximation.

EXERCISE 1 (a) Find the Taylor polynomial $P_{2,(a,b)}(x, y)$ for

$$f(x, y) = \frac{1}{2}y^2 + x - \frac{1}{3}x^3$$

at the point $(a, b) = (1, 0)$, by calculating the appropriate partial derivatives.

(b) Verify your results by letting $u = x - 1$, $v = y$ and writing

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

Expand and neglect powers higher than 2 and then convert back to x and y. This type of algebraic derivation can only be done for a polynomial function.

We now ask: How large is the error if we use the approximation

$$f(x, y) \approx P_{2,(a,b)}(x, y)?$$

To answer this question, we need to extend Taylor's Theorem to functions of two variables $f(x, y)$.

8.2 Taylor's Formula with Second Degree Remainder

Review of the 1-D case

THEOREM 1 If $f''(x)$ exists on $[a, x]$, then there exists a number c between a and x such that

$$f(x) = f(a) + f'(a)(x-a) + R_{1,a}(x) \tag{8.2}$$

where

$$R_{1,a}(x) = \frac{1}{2} f''(c)(x-a)^2 \tag{8.3}$$

On recalling that

$$L_a(x) = f(a) + f'(a)(x-a) \tag{8.4}$$

we see that the term $R_{1,a}(x)$ represents the error in using the linear approximation. Keep in mind that you can't evaluate this expression, because you don't know the value of c. We only know that c lies between a and x. However, this formula is useful because it gives a way of finding an upper bound for the error.

If f has a continuous second derivative on an interval $[a-\delta, a+\delta]$ centered on a, then f'' is bounded on this interval. That is, there exists a number B such that

$$|f''(x)| \leq B, \quad \text{for all } x \in [a-\delta, a+\delta]$$

By equations (8.2)-(8.4),

$$\begin{aligned} |f(x) - L_a(x)| &= \left|R_{1,a}(x)\right| \\ &= \left|\frac{1}{2} f''(c)(x-a)^2\right| \\ &= \frac{1}{2}|f''(c)|(x-a)^2 \\ &\leq \frac{1}{2} B(x-a)^2 \end{aligned}$$

for all $x \in [a-\delta, a+\delta]$. Knowing $f''(x)$, you can find a value for B.

The 2-D Case

In order to generalize Taylor's formula to the case of $f(x, y)$, observe that $R_{1,a}(x)$ in equation (8.3) has the same form as the second derivative term in $P_{2,a}(x)$, except that f'' is evaluated at c instead of at a. Knowing the form of $P_{2,(a,b)}(x, y)$ leads us to Taylor's Theorem for a function of two variables.

THEOREM 2 **(Taylor's Theorem)**

If $f(x,y) \in C^2$ in some neighborhood $N(a,b)$ of (a,b), then for all $(x,y) \in N(a,b)$ there exists a point (c,d) on the line segment joining (a,b) and (x,y) such that

$$f(x,y) = f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b) + R_{1,(a,b)}(x,y)$$

where

$$R_{1,(a,b)}(x,y) = \frac{1}{2}\Big[f_{xx}(c,d)(x-a)^2 + 2f_{xy}(c,d)(x-a)(y-b) + f_{yy}(c,d)(y-b)^2\Big]$$

Proof: The idea is to reduce the given function f of two variables to a function g of one variable, by considering only points on the line segment joining (a,b) and (x,y).

We parameterize the line segment L from (a,b) to (x,y) by

$$L(t) = (a + t(x-a), b + t(y-b)), \qquad 0 \le t \le 1$$

For simplicity write $h = x - a$ and $k = y - b$. Then $x - a = h$, $y - b = k$, and Taylor's formula assumes the form

$$f(x,y) = f(a,b) + f_x(a,b)h + f_y(a,b)k + R_{1,(a,b)}(x,y)$$

where

$$R_{1,(a,b)}(x,y) = \frac{1}{2}\Big[f_{xx}(c,d)h^2 + 2f_{xy}(c,d)hk + f_{yy}(c,d)k^2\Big]$$

Define g by

$$g(t) = f(L(t)), \qquad 0 \le t \le 1 \tag{8.5}$$

Since f has continuous second partials by hypothesis, we can apply the Chain Rule to conclude that g' and g'' are continuous and are given by

$$g'(t) = f_x(L(t))h + f_y(L(t))k \tag{8.6}$$

$$g''(t) = f_{xx}(L(t))h^2 + 2f_{xy}(L(t))hk + f_{yy}(L(t))k^2 \tag{8.7}$$

for $0 \le t \le 1$.

Since g'' is continuous on the interval $[0,1]$, Taylor's formula may be applied to g on this interval. That is, we can set $x = 1$ and $a = 0$ in equations (8.2) and (8.3). It follows that there exists a number $\tilde{c}$, with $0 < \tilde{c} < 1$, such that

$$g(1) = g(0) + g'(0) + \frac{1}{2}g''(\tilde{c}) \tag{8.8}$$

Each term in this equation can be calculated using equations (8.5)-(8.7), giving

$$\begin{aligned} g(1) &= f((a,b) + [(x,y) - (a,b)]) = f(x,y) \\ g(0) &= f(a,b), \text{ and} \\ g'(0) &= f_x(a,b)h + f_y(a,b)k \end{aligned}$$

In addition, if we let $(c,d) = L(\tilde{c})$, then

$$\frac{1}{2}g''(\tilde{c}) = R_{1,(a,b)}(x,y)$$

and equation (8.8) becomes precisely the modified version of Taylor's formula. □

REMARK

Like the one variable case, Taylor's Theorem for $f(x, y)$ is an existence theorem. That is, it only tells us that the point (c, d) exists, but not how to find it.

Here is an example to show how Taylor's formula can be used to estimate the error when using the linear approximation formula.

EXAMPLE 1 If $x \geq 0$ and $y \geq 0$ show that

$$\sqrt{1 + x + 2y} \approx 1 + \frac{1}{2}x + y$$

with

$$|R_{1,(0,0)}(x, y)| \leq \frac{3}{4}(x^2 + y^2)$$

Solution: By differentiating $f(x, y) = \sqrt{1 + x + 2y}$, we obtain

$$L_{(0,0)}(x, y) = 1 + \tfrac{1}{2}x + y$$

and

$$f_{xx} = \frac{-1}{4(1 + x + 2y)^{3/2}}, \quad f_{xy} = \frac{-1}{2(1 + x + 2y)^{3/2}}, \quad f_{yy} = \frac{-1}{(1 + x + 2y)^{3/2}}$$

For $x \geq 0$ and $y \geq 0$, f has continuous second partial derivatives so we can apply Taylor's Theorem to get that there exists a point (c, d) on the line segment from (x, y) to $(0, 0)$ such that

$$\left|R_{1,(0,0)}(x, y)\right| = \left|\frac{1}{2}[f_{xx}(c, d)(x - 0)^2 + 2f_{xy}(c, d)(x - 0)(y - 0) + f_{yy}(c, d)(y - 0)^2]\right|$$

Since we can not find (c, d), we want to find an upper bound for this function. Applying the triangle inequality gives

$$\left|R_{1,(0,0)}(x, y)\right| \leq \frac{1}{2}\left[\left|f_{xx}(c, d)\right| x^2 + 2\left|f_{xy}(c, d)\right| |x|\,|y| + \left|f_{yy}(c, d)\right| y^2\right] \tag{8.9}$$

Thus, to find our upper bound for the error, we just need to find upper bounds for $|f_{xx}(c, d)|$, $|f_{xy}(c, d)|$, and $|f_{yy}(c, d)|$. Since (c, d) lies on the line segment from (x, y) to $(0, 0)$, and $x \geq 0$ and $y \geq 0$ we get that $c \geq 0$ and $d \geq 0$. Consequently, $1 + c + 2d \geq 1$. Therefore,

$$\begin{aligned}
|f_{xx}(c, d)| &= \left|\frac{-1}{4(1 + c + 2d)^{3/2}}\right| \leq \frac{1}{4} \\
|f_{xy}(c, d)| &= \left|\frac{-1}{2(1 + c + 2d)^{3/2}}\right| \leq \frac{1}{2} \\
|f_{yy}(c, d)| &= \left|\frac{-1}{(1 + c + 2d)^{3/2}}\right| \leq 1
\end{aligned}$$

Substituting these into (8.9) we get

$$\begin{aligned}
\left|R_{1,(0,0)}\right| &\leq \frac{1}{2}\left[\frac{1}{4}x^2 + 2\frac{1}{2}|x|\,|y| + 1y^2\right] \\
&\leq \frac{1}{2}\left[\frac{1}{4}x^2 + \frac{1}{2}(x^2+y^2) + y^2\right], \qquad \text{since } 2|x||y| \leq x^2+y^2 \\
&= \frac{3}{8}x^2 + \frac{3}{4}y^2
\end{aligned}$$

Using the fact that $\frac{3}{8}x^2 \leq \frac{3}{4}x^2$ gives

$$\left|\sqrt{1+x+2y} - \left(1 + \frac{1}{2}x + y\right)\right| \leq \frac{3}{4}(x^2+y^2)$$

as required.

EXERCISE 1 Let $f(x,y) = e^{-2x+y}$. Use Taylor's Theorem to show that the error in the linear approximation $L_{(1,1)}(x,y)$ is at most $6e[(x-1)^2 + (y-1)^2]$ if $0 \leq x \leq 1$ and $0 \leq y \leq 1$.

REMARK

The most important thing about the error term $R_{1,(a,b)}(x,y)$ is not its explicit form, but rather its dependence on the magnitude of the displacement $\|(x,y)-(a,b)\|$. We state the result as a Corollary.

COROLLARY 3 If $f(x,y) \in C^2$ in some closed neighborhood $N(a,b)$ of (a,b), then there exists a positive constant M such that

$$|R_{1,(a,b)}(x,y)| \leq M\|(x,y)-(a,b)\|^2, \qquad \text{for all } (x,y) \in N(a,b)$$

8.3 Generalizations

In order to define the Taylor polynomial $P_{k,(a,b)}(x,y)$ of degree k, in a concise manner, we introduce the differential operator

$$(x-a)D_1 + (y-b)D_2$$

where $D_1 = \frac{\partial}{\partial x}$ and $D_2 = \frac{\partial}{\partial y}$ are the partial differential operators. Then, we formally write

$$[(x-a)D_1 + (y-b)D_2]^2 = (x-a)^2D_1^2 + 2(x-a)(y-b)D_1D_2 + (y-b)^2D_2^2$$

Note that $D_1^2 = D_1 D_1$. This means apply D_1 twice, i.e. take the second partial derivative with respect to the first variable.

In terms of this notation, the first degree Taylor polynomial $P_{1,(a,b)}(x,y)$ (which is the linear approximation $L_{(a,b)}(x,y)$) is written as

$$P_{1,(a,b)}(x,y) = f(a,b) + [(x-a)D_1 + (y-b)D_2]f(a,b)$$

and the second degree Taylor polynomial is written as

$$P_{2,(a,b)}(x,y) = P_{1,(a,b)}(x,y) + \frac{1}{2!}\left[(x-a)D_1 + (y-b)D_2\right]^2 f(a,b)$$

For $k = 2, 3, \ldots$ we recursively define the kth degree Taylor polynomial by

$$P_{k,(a,b)}(x,y) = P_{k-1,(a,b)}(x,y) + \frac{1}{k!}\left[(x-a)D_1 + (y-b)D_2\right]^k f(a,b)$$

The expression $[(x-a)D_1 + (y-b)D_2]^k$ is expanded using the Binomial Theorem.

EXERCISE 1 Write out $P_{3,(a,b)}(x,y)$ explicitly using subscript notation.

We now see that all of the results we had generalize in the expected way for all values of k.

THEOREM 1 **Taylor's Theorem of order k**
If $f(x,y) \in C^{k+1}$ at each point on the line segment joining (a,b) and (x,y), then there exists a point (c,d) on the line segment between (a,b) and (x,y) such that

$$f(x,y) = P_{k,(a,b)}(x,y) + R_{k,(a,b)}(x,y)$$

where

$$R_{k,(a,b)}(x,y) = \frac{1}{(k+1)!}\left[(x-a)D_1 + (y-b)D_2\right]^{k+1} f(c,d)$$

COROLLARY 2 If $f(x,y) \in C^{k+1}$ in some closed neighborhood $N(a,b)$ of (a,b), then there exists a constant $M > 0$ such that

$$|f(x,y) - P_{k,(a,b)}(x,y)| \leq M\,\|(x,y)-(a,b)\|^{k+1}$$

for all $(x,y) \in N(a,b)$.

COROLLARY 3 If $f(x,y) \in C^{k+1}$ in some neighborhood of (a,b), then

$$\lim_{(x,y)\to(a,b)} \frac{|f(x,y) - P_{k,(a,b)}(x,y)|}{\|(x,y)-(a,b)\|^k} = 0$$

The final stage in the process of generalization is to consider functions of n variables $f(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^n$. One has simply to replace the differential operator

$$[(x-a)D_1 + (y-b)D_2]$$

by

$$[(x_1-a_1)D_1 + \cdots + (x_n-a_n)D_n]$$

Letting $\nabla = (D_1, \ldots, D_n)$, we can be write this concisely in vector notation as

$$\left[(\mathbf{x}-\mathbf{a}) \cdot \nabla\right]$$

Chapter 8 Problem Set

1. Let $f(x,y) = e^{3x-2y}$.
   (a) Calculate the gradient vector and the Hessian matrix of f at $(a,b) = (2,3)$.
   (b) Write down the linearization $L_{(2,3)}(x,y)$ and the Taylor polynomial $P_{2,(2,3)}(x,y)$ of f.
   (c) Show that the gradient vector of f has the same direction at each point. What conclusion can you draw about the level curves of f?
2. Find the Taylor polynomial $P_{2,(a,b)}(x,y)$ for each function.
   (a) $f(x,y) = \ln(x+e^y)$, $(a,b) = (1,0)$
   (b) $f(x,y) = xe^{x-y}$, $(a,b) = (1,1)$
3. Let $f(x,y) = (x-y)\sin(x+y)$. Find the Taylor polynomial $P_{2,(a,b)}(x,y)$ of f at (π,π).
4. Use the second degree Taylor polynomial to derive the approximation $\ln(\sin^2 x + \cos^2 y) \approx x^2 - y^2$ for (x,y) sufficiently close to (0,0).
5. (a) Use the second degree Taylor polynomial to derive the approximation $(1+x)^y \approx 1 + xy$ for (x,y) sufficiently close to $(0,0)$.
   (b) Test the accuracy of the approximation in (a) with your calculator by making a table of values (3 cases). Give the percentage error in the approximations.
6. Consider the approximation

$$\ln(x+2y) \approx (x-3) + 2(y+1)$$

   for (x,y) sufficiently close to $(3,-1)$. Prove that if $x \geq 3$ and $y \geq -1$, the error satisfies

$$|\text{error}| \leq \frac{7}{2}\left[(x-3)^2 + (y+1)^2\right]$$

左圖
沒考

7. Find a function $f(x,y)$ such that $Hf(x,y) = \begin{bmatrix} 1 & 2 \\ 2 & -3 \end{bmatrix}$ for all $(x,y) \in \mathbb{R}^2$, $\nabla f(1,0) = (-2,5)$ and $f(1,0) = 7$. Is there more than one such f?
8. Suppose that $f(x,y)$ has continuous second partial derivatives which satisfy $|f_{xx}| \leq M$, $|f_{xy}| \leq M$, $|f_{yy}| \leq M$ for all $(x,y) \in N = \{(x,y) \mid (x-a)^2 + (y-b)^2 \leq r^2\}$, where M is a constant. Let $L_{(a,b)}(x,y)$ be the linear approximation of f at (a,b). Prove that

$$|f(x,y) - L_{(a,b)}(x,y)| \leq M[(x-a)^2 + (y-b)^2]$$

   for all $(x,y) \in N$. This gives an upper bound for the error in the linear approximation.
9. Let $f(x,y) = e^{x-4y}$. Use Taylor's Theorem to show that the error in the linear approximation $L_{(1,1)}(x,y)$ is at most $\frac{e}{2}[5(x-1)^2 + 20(y-1)^2]$ if $0 \leq x \leq 1$ and $0 \leq y \leq 1$.
10. Let $f(x,y) = \ln(1+x+2y)$. Use Taylor's Theorem to show that for $x \geq 0$, $y \geq 0$ we have

$$|R_{1,(0,0)}(x,y)| \leq 3(x^2+y^2)$$

11. Let $f(x,y) = \frac{1}{xy}$ for $x > 0$ and $y > 0$. Use Taylor's Theorem to show that if $x > 1$ and $y > 1$, then

$$|f(x,y) - L_{(1,1)}(x,y)| \leq \frac{3}{2}[(x-1)^2 + (y-1)^2]$$

12. Consider a function f defined by $f(x,y) = 2x^2 + 3y^2$, and let $(a,b) \in \mathbb{R}^2$ be arbitrary. Prove that $f(x,y) \geq L_{(a,b)}(x,y)$, for all $(x,y) \in \mathbb{R}^2$.
   Comment: Since $z = L_{(a,b)}(x,y)$ is the equation of the tangent plane to the surface $z = f(x,y)$ at (a,b), this shows that the surface lies above each of its tangent planes.
13. * Suppose that $f(x,y)$ has continuous second partial derivatives on the rectangle $a \leq x \leq b, c \leq y \leq d$. Use Taylor's formula to prove that

$$\frac{d}{dx}\int_c^d f(x,y)dy = \int_c^d \frac{\partial f(x,y)}{\partial x}dy$$

   for all x which satisfy $a < x < b$.
   Hint: Let $g(x) = \displaystyle\int_c^d f(x,y)\,dy$, and use the definition of the derivative to calculate $g'(x)$.

1. Let $f(x,y) = e^{3x-2y}$.
   (a) Calculate the gradient vector and the Hessian matrix of f at $(a,b) = (2,3)$.
   (b) Write down the linearization $L_{(2,3)}(x,y)$ and the Taylor polynomial $P_{2,(2,3)}(x,y)$ of f.
   (c) Show that the gradient vector of f has the same direction at each point. What conclusion can you draw about the level curves of f?

(a) $\nabla f(x,y) = (3e^{3x-2y}, -2e^{3x-2y})$

$\nabla f(2,3) = (3,-2)$

$f_{xx} = 9e^{3x-2y}$ $\quad f_{xy} = -6e^{3x-2y}$

$f_{yy} = 4e^{3x-2y}$ $\quad Hf(x,y) = \begin{bmatrix} 9 & -6 \\ -6 & 4 \end{bmatrix}$

(b) $L_{(2,3)}(x,y) = 1 + 3(x-2) - 2(x-3)$

$P_{2,(2,3)}(x,y) = 1 + 3(x-2) - 2(x-3)$
$+ \frac{1}{2}[9(x-2)^2 - 12(x-2)(y-3) + 4(y-3)^2]$

(c) $\nabla f = (3e^{3x-2y}, -2e^{3x-2y}) = e^{3x-2y}(3,-2)$

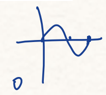
$\therefore$ gradient vector is always in the direction of $(3,-2)$ since $e^{3x-2y} > 0 \ \forall (x,y)$

conclusion:

by orthogonality theorem, ∇f is always orthogonal to the level curves of f.

Since gradient vectors, are all in same direction for any (x,y) so level curves are of the form parallel straight lines.

3. Let $f(x,y) = (x-y)\sin(x+y)$. Find the Taylor polynomial $P_{2,(a,b)}(x,y)$ of f at (π,π).



$f(\pi,\pi) = 0$

$f_x = \sin(x+y) + (x-y)\cos(x+y)$ $\quad 0$

$f_y = -\sin(x+y) + (x-y)\cos(x+y)$ $\quad 0$

$f_{xx} = \cos(x+y) + \cos(x+y) - (x-y)\sin(x+y)$

$f_{xy} = \cos(x+y) - \cos(x+y) - (x-y)\sin(x+y)$

$f_{yy} = -\cos(x+y) - \cos(x+y) - (x-y)\sin(x+y)$

$Hf(\pi,\pi) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$

$P_{2,(a,b)}(x,y) = \frac{1}{2}[2(x-\pi)^2 - 2(y-\pi)^2]$

$= (x-\pi)^2 - (y-\pi)^2$

2. Find the Taylor polynomial $P_{2,(a,b)}(x,y)$ for each function.
   (a) $f(x,y) = \ln(x + e^y)$, $(a,b) = (1,0)$
   (b) $f(x,y) = xe^{x-y}$, $(a,b) = (1,1)$

(a) $f(1,0) = 0$

$f_x = (x+e^y)^{-1}$ $\quad f_x(1,0) = 1$

$f_y = e^y(x+e^y)^{-1}$ $\quad f_y(1,0) = 1$

$f_{xx} = -(x+e^y)^{-2}.$

$f_{xy} = -(x+e^y)^{-2} \cdot e^y$ $\quad Hf(1,0) = \begin{bmatrix} -1 & -1 \\ -1 & 0 \end{bmatrix}$

$f_{yy} = e^y(x+e^y)^{-1} + e^y \cdot (-(x+e^y)^{-2} \cdot e^y) = e^y(x+e^y)^{-1} - e^{2y}(x+e^y)^{-2}$

$P_{2(1,0)}(x,y) = (x-1) + (y-1) + \frac{1}{2}[-1(x-1)^2 - 2(x-1)(y-1)]$

(b) $f(1,1) = 1$

$f_x = e^{x-y} + xe^{x-y}$ $\quad f_x(1,0) = 1+1 = 2$

$f_y = -xe^{x-y}$ $\quad f_y(1,0) = -1$

$f_{xx} = 2e^{x-y} + xe^{x-y}$

$f_{xy} = -e^{x-y} - xe^{x-y}$ $\quad Hf(1,1) = \begin{bmatrix} 3 & -2 \\ -2 & 1 \end{bmatrix}$

$f_{yy} = xe^{x-y}$

$P_{2(1,1)}(x,y) = 1 + 2(x-1) - y + \frac{1}{2}[3(x-1)^2 - 4(x-1)y + y^2]$

4. Use the second degree Taylor polynomial to derive the approximation $\ln(\sin^2 x + \cos^2 y) \approx x^2 - y^2$ for (x, y) sufficiently close to $(0,0)$.

let $f(x,y) = \ln(\sin^2 x + \cos^2 y)$

$f(0,0) = 0$

$f_x = 2\cos x \sin x \cdot (\sin^2 x + \cos^2 y)^{-1}$

$f_y = -2\sin y \cos y \cdot (\sin^2 x + \cos^2 y)^{-1}$

$f_{xx} = -\sin x(\sin^2 x + \cos^2 y)^{-1} + 2\cos x(-(\sin^2 x + \cos^2 y)^{-2} \cdot 2\sin x \cos x)$

$f_{xy} = 2\cos x \cdot (-(\sin^2 x + \cos^2 y)^{-2} \cdot 2\cos y(-\sin y))$ $\quad 2 \cdot 2$

$f_{yy} = -2\cos y(\sin^2 x + \cos^2 y)^{-1} - 2\sin y(-(\sin^2 x + \cos^2 y)^{-2} \cdot 2\cos y(-\sin y))$

$Hf(0,0) = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}$

$P_{2(0,0)}(x,y) = \frac{1}{2}(2x^2 - 2y^2) = x^2 - y^2.$

5. (a) Use the second degree Taylor polynomial to derive the approximation $(1 + x)^y \approx 1 + xy$ for (x, y) sufficiently close to $(0, 0)$.
   (b) Test the accuracy of the approximation in (a) with your calculator by making a table of values (3 cases). Give the percentage error in the approximations.

(a) $f(x,y) = (1+x)^y$ $\quad 1$

$\frac{d}{dx}(a^x) = a^x \ln a$

$f_x = y(1+x)^{y-1}$ $\quad 0$

$f_y = (1+x)^y \ln \cancel{y}(1+x)$ $\quad 0$

$f_{xx} = y \cdot (y-1)(1+x)^{y-2}$

$f_{xy} = (1+x)^{y-1} + (1+x)^{y-1}\ln(1+x)$

$f_{yy} = y(1+x)^{y-1}\ln y + (1+x)^y \cdot \frac{1}{y}$ ~~← 可直接得0)~~

算的有拉

$Hf(0,0) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

$y = 1 + \frac{1}{2}(2xy) = 1 + xy.$

(b) 略

6. Consider the approximation

$$\ln(x + 2y) \approx (x - 3) + 2(y + 1)$$

for (x, y) sufficiently close to $(3, -1)$. Prove that if $x \geq 3$ and $y \geq -1$, the error satisfies

$$|\text{error}| \leq \frac{7}{2}\left[(x - 3)^2 + (y + 1)^2\right]$$

let $f(x,y) = \ln(x+2y)$

$f(3,-1) = 0$

$f_x = (x+2y)^{-1}$ $\quad 1$

$f_y = 2(x+2y)^{-1}$ $\quad 2$

$f_{xx} = -(x+2y)^{-2}.$ $\quad -1$

$f_{xy} = -2(x+2y)^{-1}$ $\quad -2$

$f_{yy} = -2(x+2y)^{-1} \cdot 2 = -4(x+2y)^{-1}$ $\quad -4$

$Hf(3,-1) = \begin{bmatrix} -1 & -2 \\ -2 & -4 \end{bmatrix}$

$R_{2(1,3)}(x,y) = \frac{1}{2}[-1(x-3)^2 - 4(x-3)(y+1) - 4(y+1)^2]$

$x \geq 3 \quad y \geq -1.$

$4(x-3)(y+1) \geq 0$

$R_{2(1,3)}(x,y) \geq -\frac{1}{2}(x-3)^2 - 2(y+1)^2$

$\leq \frac{1}{2}(x+3)^2 + 2(y+1)^2$

$\downarrow \qquad \downarrow$

$< \frac{7}{2} \qquad < \frac{7}{2}$

$\leq \frac{7}{2} \ldots$

7. Find a function $f(x,y)$ such that $Hf(x,y) = \begin{bmatrix} 1 & 2 \\ 2 & -3 \end{bmatrix}$ for all $(x,y) \in \mathbb{R}^2$, $\nabla f(1,0) = (-2,5)$ and $f(1,0) = 7$. Is there more than one such f?

$$\begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & -3 \end{bmatrix}.$$

$f_x(1,0) = -2 \quad f_y(1,0) = 5$

$f(1,0) = 7.$

8. Suppose that $f(x,y)$ has continuous second partial derivatives which satisfy $|f_{xx}| \le M$, $|f_{xy}| \le M$, $|f_{yy}| \le$

M for all $(x,y) \in N = \{(x,y) \mid (x-a)^2 + (y-b)^2 \le r^2\}$, where M is a constant. Let $L_{(a,b)}(x,y)$ be the linear approximation of f at (a,b). Prove that

$$|f(x,y) - L_{(a,b)}(x,y)| \le M[(x-a)^2 + (y-b)^2]$$

for all $(x,y) \in N$. This gives an upper bound for the error in the linear approximation.

Chapter 9

Critical Points

Recall from single variable calculus that if $x = a$ is a local extremum of $f(x)$, then either $f'(a) = 0$ or $f'(a)$ does not exist. Such points are of interest and are called critical points of f. But, recall that a critical point is not necessarily a local extremum. For example, $f(x) = x^3$ at $x = 0$.

In this chapter, we extend these ideas to functions $f(x, y)$. The second degree Taylor polynomial will be used to generalize the second derivative test for local extrema. These ideas will be applied to optimization problems in Chapter 10.

9.1 Local Extrema and Critical Points

We begin with the definitions of local extrema.

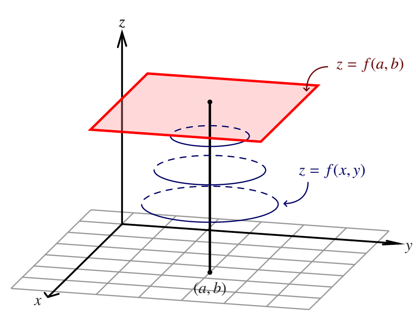
DEFINITION

Local Maximum and Minimum

A point (a, b) is a **local maximum point of** f if $f(x, y) \le f(a, b)$ for all (x, y) in some neighborhood of (a, b).

A point (a, b) is a **local minimum point of** f if $f(x, y) \ge f(a, b)$ for all (x, y) in some neighborhood of (a, b).

Thinking geometrically, if (a, b) is a local maximum/minimum point of f and f has continuous partial derivatives, then (a, b) is a local maximum/minimum point of the cross-sections $f(x, b)$ and $f(a, y)$. Thus, (a, b) is a critical point of both of these cross-sections and so both partial derivatives of f will be zero and the tangent plane will be horizontal.



THEOREM 1

If (a, b) is a local maximum or minimum point of f, then

$$f_x(a, b) = 0 = f_y(a, b)$$

or at least one of f_x or f_y does not exist at (a, b).

Proof: Consider the function g defined by $g(x) = f(x, b)$. If (a, b) is a local maximum/minimum point of f, then $x = a$ is a local maximum/minimum point of g, and hence either $g'(a) = 0$ or $g'(a)$ does not exist. Thus it follows that either $f_x(a, b) = 0$ or $f_x(a, b)$ does not exist. A similar argument gives $f_y(a, b) = 0$ or $f_y(a, b)$ does not exist. $\square$

DEFINITION
Critical Point

A point (a, b) in the domain of $f(x, y)$ is called a **critical point** of f if $\dfrac{\partial f}{\partial x}(a, b) = 0$ or $\dfrac{\partial f}{\partial x}(a, b)$ does not exist, and $\dfrac{\partial f}{\partial y}(a, b) = 0$ or $\dfrac{\partial f}{\partial y}(a, b)$ does not exist.

EXAMPLE 1

Find the critical points of the following functions and determine if they are local maximum points or local minimum points.

$$f(x, y) = x^2 + y^2, \qquad g(x, y) = -x^2 - y^2, \qquad h(x, y) = x^2 - y^2$$

Solution: For f, we see that $f_x(x, y) = 2x$ and $f_y(x, y) = 2y$ so $(0, 0)$ is the only critical point of f. Observe that

$$f(x, y) = x^2 + y^2 > 0 = f(0, 0), \qquad \text{for all } (x, y) \neq (0, 0)$$

so $(0, 0)$ is a local minimum point of f.

For g, we have $g_x(x, y) = -2x$ and $g_y(x, y) = -2y$ so $(0, 0)$ is the only critical point of g and

$$g(x, y) = -x^2 - y^2 < 0 = g(0, 0), \qquad \text{for all } (x, y) \neq (0, 0)$$

so $(0, 0)$ is a local maximum point of g.

For h, we have $h_x(x, y) = 2x$ and $h_y(x, y) = -2y$ so $(0, 0)$ is the only critical point of h, but we have $h(x, 0) > h(0, 0)$ for any value of x and $h(0, y) < h(0, 0)$ for any value of y, so $(0, 0)$ is neither a local maximum point nor a local minimum point.

Our solutions for f and g make a lot of sense when we realize that $z = f(x, y)$ is a paraboloid facing up and $z = g(x, y)$ is a paraboloid facing down. Also, we see that $(0, 0)$ is the point at the center of the saddle for the saddle surface $z = h(x, y)$ hence it should not be a local minimum or a local maximum. This motivates the following definition.

DEFINITION
Saddle Point

A critical point (a, b) of $f(x, y)$ is called a **saddle point** of f if in every neighborhood of (a, b) there exist points (x_1, y_1) and (x_2, y_2) such that

$$f(x_1, y_1) > f(a, b) \text{ and } f(x_2, y_2) < f(a, b)$$

The problem that we are faced with has two parts.

(1) Given $f(x, y)$, find all critical points of f.
(2) Determine whether the critical points are local maxima, minima or saddle points.

We now illustrate (1) with an example. (2) is discussed in Section 9.2.

EXAMPLE 2

Find all critical points of $f(x, y) = x^2y + 3xy^2 + xy$.

Solution: Differentiate and simplify, to obtain

$$\frac{\partial f}{\partial x}(x, y) = y(2x + 3y + 1), \qquad \frac{\partial f}{\partial y}(x, y) = x(x + 6y + 1)$$

In this type of problem it is helpful to take out common factors in the expressions. To find the critical points of f we must solve the system of two equations

$$y(2x + 3y + 1) = 0 \tag{9.1}$$
$$x(x + 6y + 1) = 0 \tag{9.2}$$

Observe that (9.1) implies that either $y = 0$ or $2x + 3y + 1 = 0$. We consider these two cases:

Case 1: $y = 0$.
Putting $y = 0$ into (9.2) we get $x(x + 1) = 0$, giving two values $x = 0$ or $x = -1$. Thus, we have critical points $(0, 0)$ and $(-1, 0)$.

Case 2: $2x + 3y + 1 = 0$.
We have $3y = -2x - 1$, so (9.2) gives

$$0 = x(x + 2(3y) + 1) = x(x + 2(-2x - 1) + 1) = -3x^2 - x = -x(3x + 1)$$

giving two values $x = 0$ and $x = -\frac{1}{3}$. To find the corresponding y values we put these into $3y = -2x - 1$ and get two more critical points $(0, -\frac{1}{3})$ and $(-\frac{1}{3}, -\frac{1}{9})$.

So, the critical points are $(0, 0)$, $\left(0, -\frac{1}{3}\right)$, $(-1, 0)$, and $\left(-\frac{1}{3}, -\frac{1}{9}\right)$.

REMARK

1) It is essential to solve equations (9.1) and (9.2) systematically, by considering all possible cases, in order to find all critical points.

2) You should be aware that we can only explicitly find the critical points for simple functions f. The equations

$$f_x(x, y) = 0, \qquad f_y(x, y) = 0$$

are a system of equations which are generally non-linear, and there is no general algorithms for solving such systems exactly. There are, however, numerical methods for finding approximate solutions, one of which is a generalization of Newton's method. If you review the one variable case, you might see how to generalize it, using the tangent plane. It's a challenge!

EXERCISE 1 Find all critical points of $f(x, y) = xye^{x-y}$.

EXERCISE 2 Find all critical points of $f(x, y) = x\cos(x + y)$.

EXERCISE 3 Give a function $f(x, y)$ with no critical points.

9.2 The Second Derivative Test

Review of the 1-D case

For a function $f(x)$ of one variable, the second degree Taylor polynomial approximation is

$$f(x) \approx f(a) + f'(a)(x - a) + \frac{1}{2}f''(a)(x - a)^2$$

for x sufficiently close to a. If $x = a$ is a critical point of f, then $f'(a) = 0$, and the approximation can be rearranged to give

$$f(x) - f(a) \approx \frac{1}{2}f''(a)(x - a)^2$$

Thus, for x sufficiently close to a, $f(x)-f(a)$ has the same sign as $f''(a)$. If $f''(a) > 0$, then $f(x) - f(a) > 0$ for x sufficiently close to a and $x = a$ is a local minimum point. If $f''(a) < 0$, then $f(x) - f(a) < 0$ for x sufficiently close to a and $x = a$ is a local maximum point. There is no conclusion if $f''(a) = 0$.

The 2-D Case

For $f(x, y) \in C^2$, the second degree Taylor polynomial approximation is

$$f(x, y) \approx P_{2,(a,b)}(x, y)$$

for (x, y) sufficiently close to (a, b). If (a, b) is a critical point of f such that

$$f_x(a, b) = 0 = f_y(a, b)$$

then the approximation can be rearranged to yield

$$f(x, y) - f(a, b) \approx \tfrac{1}{2}\Big[f_{xx}(a, b)(x-a)^2 + 2f_{xy}(a, b)(x-a)(y-b) + f_{yy}(a, b)(y-b)^2\Big] \quad (9.3)$$

for (x, y) sufficiently close to (a, b). The sign of the expression on the right will determine the sign of $f(x, y) - f(a, b)$, and hence whether (a, b) is a local maximum, local minimum or saddle point.

The expression on the right is called a **quadratic form**, and at this stage it is necessary to discuss some properties of these objects.

Quadratic Forms

DEFINITION
Quadratic Form

A function Q of the form

$$Q(u, v) = a_{11}u^2 + 2a_{12}uv + a_{22}v^2$$

where a_{11}, a_{12} and a_{22} are constants, is called a **quadratic form** on $\mathbb{R}^2$.

It is important to observe that one can use matrix notation and write

$$Q(u, v) = [u \quad v]\begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix}$$

so that a quadratic form on $\mathbb{R}^2$ is determined by a 2×2 symmetric matrix.

We classify quadratic forms on $\mathbb{R}^2$ in the following way:

(1) If $Q(u, v) > 0$ for all $(u, v) \neq (0, 0)$, then $Q(u, v)$ is said to be **positive definite**.
(2) If $Q(u, v) < 0$, for all $(u, v) \neq (0, 0)$, then $Q(u, v)$ is said to be **negative definite**.
(3) If $Q(u, v) < 0$ for some (u, v) and $Q(u, v) > 0$ for some other (u, v), then $Q(u, v)$ is said to be **indefinite**.
(4) If $Q(u, v)$ does not satisfy any of 1) – 3), then $Q(u, v)$ is said to be **semidefinite**.

These terms are also used to describe the corresponding symmetric matrices.

EXAMPLE 1

$A = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$ is positive definite, since $Q(u, v) = 2u^2 + 3v^2 > 0$, for all $(u, v) \neq (0, 0)$.

$B = \begin{bmatrix} 2 & 0 \\ 0 & -3 \end{bmatrix}$ is indefinite, since $Q(u, v) = 2u^2 - 3v^2$, and $Q(u, 0) = 2u^2 > 0$ for $u \neq 0$, and $Q(0, v) = -3v^2 < 0$ for $v \neq 0$.

$C = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$ is semidefinite, since $Q(u, v) = 2u^2 \geq 0$ for all (u, v), and $Q(0, v) = 0$ for all v.

REMARK

Semidefinite quadratic forms may be split into two classes, positive semidefinite and negative semidefinite. The matrix C above would be classified as positive semidefinite.

If A is not a diagonal matrix, the nature of A (or of $Q(u, v)$) is not immediately obvious. For example, even if all entries of A are positive, it does not follow that A is a positive definite matrix.

EXAMPLE 2

Classify the symmetric matrix $A = \begin{bmatrix} 1 & 3 \\ 3 & 2 \end{bmatrix}$.

Solution: The associated quadratic form is

$$Q(u, v) = u^2 + 6uv + 2v^2$$

Complete the square, obtaining

$$Q(u, v) = (u + 3v)^2 - 7v^2$$

It is now clear by inspection that A is indefinite, since

$$Q(u, 0) = u^2 > 0, \quad \text{for } u \neq 0$$

and

$$Q(-3v, v) = -7v^2 < 0, \quad \text{for } v \neq 0$$

Having introduced quadratic forms, we return to equation (9.3). Let

$$u = x - a, \qquad v = y - b$$

so that

$$f(x,y) - f(a,b) \approx \frac{1}{2}\Big[f_{xx}(a,b)u^2 + 2f_{xy}(a,b)uv + f_{yy}(a,b)v^2\Big]$$

The matrix of the quadratic form on the right is the Hessian matrix of f at (a,b):

$$Hf(a,b) = \begin{bmatrix} f_{xx}(a,b) & f_{xy}(a,b) \\ f_{xy}(a,b) & f_{yy}(a,b) \end{bmatrix}$$

It is thus plausible that if $Hf(a,b)$ is positive definite, then

$$f(x,y) - f(a,b) > 0$$

for all $(u,v) \neq (0,0)$ i.e. for all $(x,y) \neq (a,b)$ (assuming, of course, that (x,y) is sufficiently close to (a,b) so that the approximation is sufficiently accurate). In other words, if $Hf(a,b)$ is positive definite, it is plausible that (a,b) is a local minimum point of f. One can give similar arguments in the cases where $Hf(a,b)$ is negative definite or indefinite, leading to the following theorem.

THEOREM 1

Second Partial Derivatives Test

Suppose that $f(x,y) \in C^2$ in some neighborhood of (a,b) and that

$$f_x(a,b) = 0 = f_y(a,b)$$

(1) If $Hf(a,b)$ is positive definite, then (a,b) is a local minimum point of f.

(2) If $Hf(a,b)$ is negative definite, then (a,b) is a local maximum point of f.

(3) If $Hf(a,b)$ is indefinite, then (a,b) is a saddle point of f.

REMARKS

(1) The argument preceding the theorem is not a proof, since it involves an approximation. One can use Taylor's formula and a continuity argument to give a proof. See Section 9.2.

(2) Note the analogy with the second derivative test for functions of one variable. The requirement $g''(a) > 0$, which implies a local minimum, is replaced by the requirement that the matrix of second partial derivatives $Hf(a,b)$ be positive definite.

To help us classify the Hessian matrix we can use the following theorem from the theory of quadratic forms.

THEOREM 2

If $Q(u,v) = a_{11}u^2 + 2a_{12}uv + a_{22}v^2$ and $D = a_{11}a_{22} - a_{12}^2$, then

(1) Q is positive definite if and only if $D > 0$ and $a_{11} > 0$

(2) Q is negative definite if and only if $D > 0$ and $a_{11} < 0$

(3) Q is indefinite if and only if $D < 0$

(4) Q is semidefinite if and only if $D = 0$

REMARK

Observe that D is the determinant of the associated symmetric matrix.

EXAMPLE 3

Find and classify all critical points of the function $f(x,y) = x^3 - 4x^2 + 4x - 4xy^2$.

Solution: To find the critical points we solve the system

$$0 = f_x(x,y) = 3x^2 - 8x + 4 - 4y^2 \tag{9.4}$$
$$0 = f_y(x,y) = -8xy \tag{9.5}$$

From (9.5) we get that $x = 0$ or $y = 0$. If $x = 0$, then (9.4) gives $0 = 4 - 4y^2$ so $y = \pm 1$. If $y = 0$, then (9.4) gives $0 = 3x^2 - 8x + 4 = (3x-2)(x-2)$. Hence, we have critical points $(0,1)$, $(0,-1)$, $(2,0)$, and $\left(\frac{2}{3}, 0\right)$.

The second partial derivatives are

$$f_{xx}(x,y) = 6x - 8, \quad f_{xy}(x,y) = -8y, \quad f_{yy}(x,y) = -8x$$

At $\left(\frac{2}{3}, 0\right)$, the Hessian matrix is $Hf\left(\frac{2}{3}, 0\right) = \begin{bmatrix} -4 & 0 \\ 0 & -\frac{16}{3} \end{bmatrix}$, which is clearly negative definite, since the corresponding quadratic form is $Q(u,v) = -4u^2 - \frac{16}{3}v^2$. Thus, by the second partial derivative test, $\left(\frac{2}{3}, 0\right)$ is a local maximum point.

At $(0,1)$, we get $Hf(0,1) = \begin{bmatrix} -8 & -8 \\ -8 & 0 \end{bmatrix}$. So, $\det Hf(0,1) = -64 < 0$. Thus $Hf(0,1)$ is indefinite, and by the second partial derivative test, $(0,1)$ is a saddle point.

Similarly, it follows that $(0,-1)$ and $(2,0)$ are saddle points.

EXERCISE 1

Fill in the details of Example 3 above.

EXERCISE 2

Find and classify all critical points of the function $f(x,y) = x^2 + 6xy + 2y^2$.

EXERCISE 3 Find and classify all critical points of the function $f(x, y) = (x^2 + y^2 - 1)y$.

REMARK

Another way of classifying the Hessian matrix is by finding its eigenvalues. In particular, a symmetric matrix is positive definite if all of its eigenvalues are positive, negative definite if all of its eigenvalues are negative, and indefinite if has both positive and negative eigenvalues.

Degenerate Critical Points

We have seen that quadratic forms (i.e. symmetric matrices) can be classified into four types: positive definite, negative definite, indefinite and semidefinite. Note that the second partial derivative test gives a conclusion in the first three cases but makes no reference to the semidefinite case. In fact, if $Hf(a, b)$ is semidefinite, the critical point (a, b) may be a local maximum point, a local minimum point or a saddle point. We justify this statement by considering the functions

$$f(x, y) = x^4 + y^4, \quad g(x, y) = x^4 - y^4, \quad h(x, y) = -x^4 - y^4$$

For each function $(0, 0)$ is the only critical point, and the Hessian matrix at $(0, 0)$ is the zero matrix, which is semidefinite. However, since

$$\begin{aligned} f(x, y) - f(0, 0) &\geq 0 \quad \text{for all } (x, y) \\ g(x, 0) - g(0, 0) &\geq 0 \quad \text{for all } x \\ g(0, y) - g(0, 0) &\leq 0 \quad \text{for all } y \\ h(x, y) - h(0, 0) &\leq 0 \quad \text{for all } (x, y) \end{aligned}$$

it follows that $(0, 0)$ is a local minimum point for f, a saddle point for g and a local maximum point for h.

If $Hf(a, b)$ is semidefinite, so that the second partial derivative test gives no conclusion, we say that the critical point (a, b) is **degenerate**. In order to classify the critical point, one has to investigate the sign of $f(x, y) - f(a, b)$ in a small neighborhood of (a, b).

EXAMPLE 4 Show that $(0, 0)$ is a degenerate critical point of $f(x, y) = 2(x - y)^2 - x^4 - y^4 + 3$ and classify it.

Solution: It is a routine matter to show that

$$\nabla f(0, 0) = (0, 0), \qquad Hf(0, 0) = \begin{bmatrix} 4 & -4 \\ -4 & 4 \end{bmatrix}$$

The quadratic form associated with the Hessian is

$$Q(u, v) = 4u^2 - 8uv + 4v^2 = 4(u - v)^2 \geq 0$$

with $Q(u, u) = 0$ for all u, hence $Hf(0, 0)$ is semidefinite. Thus, $(0, 0)$ is a degenerate critical point. In order to classify it, consider

$$f(x, y) - f(0, 0) = 2(x - y)^2 - x^4 - y^4$$

Observe that

$$f(x, x) - f(0, 0) = -2x^4 < 0 \quad \text{for all } x \neq 0$$

and

$$f(x, 0) - f(0, 0) = 2x^2 - x^4 = x^2(2 - x^2) > 0$$

for all x which satisfy $0 < x^2 < 2$. So, in any sufficiently small neighborhood of $(0, 0)$, $f(x, y) - f(0, 0)$ assumes positive and negative values. Hence, $(0, 0)$ is a saddle point.

Generalizations

The definitions of local maximum point, local minimum point and critical point can be generalized in the obvious way to functions f of n variables. The Hessian matrix of f at $\mathbf{a}$ is the $n \times n$ symmetric matrix given by

$$Hf(\mathbf{a}) = \left[\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a})\right]$$

where $i, j = 1, 2, \ldots, n$. The Hessian matrix can be classified as positive definite, negative definite, indefinite or semidefinite by considering the associated quadratic form in $\mathbb{R}^n$:

$$Q(\mathbf{u}) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a, b) u_i u_j$$

as in $\mathbb{R}^2$. The second derivative test as stated in $\mathbb{R}^2$ now holds in $\mathbb{R}^n$. It can be justified heuristically by using the second degree Taylor polynomial approximation,

$$f(\mathbf{x}) \approx P_{2,\mathbf{a}}(\mathbf{x})$$

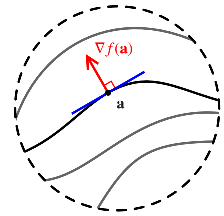
which leads to

$$f(\mathbf{x}) - f(\mathbf{a}) \approx \frac{1}{2!} \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a, b)(x_i - a_i)(x_j - a_j)$$

generalizing equation (9.3).

Level Curves Near a Critical Point

Consider a function $f \in C^2$. In Section 7.2 we discussed the fact that if $\nabla f(a,b) \neq (0,0)$, then the level curve of f through (a,b) is a smooth curve (at least sufficiently close to (a,b)). Also, by continuity, $\nabla f(x,y) \neq (0,0)$ for all (x,y) in some neighborhood of (a,b). Thus, if $\nabla f(a,b) \neq (0,0)$, there will be some neighborhood of (a,b) in which the level curves of f are smooth non-intersecting curves.



Level curves $f(x,y) = k$ near a, b when $\nabla f(a,b) \neq (0,0)$ (actual shape is not significant)

A point at which $\nabla f(a,b) \neq (0,0)$ is called a **regular point** of f.

Assume that f has continuous second partial derivatives, and approximate f by its Taylor polynomial $P_{2,(a,b)}(x,y)$, calculated at the critical point:

$$f(x,y) \approx f(a,b) + \frac{1}{2}\Big[f_{xx}(a,b)(x-a)^2 + 2f_{xy}(a,b)(x-a)(y-b) + f_{yy}(a,b)(y-b)^2\Big] \quad (9.6)$$

The constant term $f(a,b)$ and the factor $\frac{1}{2}$ in equation (9.6) do not make a significant difference to the shape of the level curves. So, it is plausible (and can be proven) that the level curves of f will be approximated by the level curves of $P_{2,(a,b)}(x,y)$ for (x,y) sufficiently close to (a,b).

Performing the translation $u = x - a$ and $v = y - b$, we get a quadratic form

$$Q(u,v) = a_{11}u^2 + 2a_{12}uv + a_{22}v^2$$

Therefore, to approximate the level curves of f near a critical point, we can sketch the level curves of the associated quadratic form $Q(u,v)$.

To sketch level curves of quadratic forms requires even more linear algebra. The possible shapes and how to sketch them is covered in Math 235.

Convex Functions

1-D Case

We say that a twice differentiable function $f(x)$ is **strictly convex** if $f''(x) > 0$ for all x and f is **convex** is $f''(x) \geq 0$ for all x. Thus the term convex means “concave up.” Convex functions have two interesting properties.

THEOREM 3 If $f(x)$ is twice continuously differentiable and strictly convex, then

(1) $f(x) > L_a(x) = f(a) + f'(x)(x-a)$ for all $x \neq a$, for any $a \in \mathbb{R}$.

(2) For $a < b$, $f(x) < f(a) + \frac{f(b)-f(a)}{b-a}(x-a)$ for $x \in (a,b)$.

Proof: (1) Follows from Taylor's Theorem: $f(x) = L_a(x) + \frac{f''(c)}{2}(x-a)^2$ where c is between a and x. Thus $R_{1,a}(x) > 0$ for $x \neq a$, giving $f(x) > L_a(x)$ for all $x \neq a$.

(2) Let $g(x) = f(x) - \left[f(a) + \frac{f(b)-f(a)}{b-a}(x-a)\right]$. Then $g(a) = g(b) = 0$ and $g''(x) = f''(x) > 0$. We must show that $g(x) < 0$ for $x \in (a,b)$. By the Mean Value Theorem $\frac{f(b)-f(a)}{b-a} = f'(c)$ for some $c \in (a,b)$. Note that $g'(x) = f'(x) - \frac{f(b)-f(a)}{b-a} = f'(x) - f'(c)$. Thus $g'(c) = 0$. Since $g''(x) > 0$ then $g'(x)$ is strictly increasing. Since $g'(c) = 0$ then $g'(x) < 0$ on $[a,c)$ and $g'(x) > 0$ on $(c,b]$. This implies that $g(x)$ is strictly decreasing on $[a,c]$ and strictly increasing on $[c,b]$. Since $g(a) = 0$ and $g(b) = 0$ we get that $g(x) < 0$ on $(a,c]$ and on $[c,b)$. Therefore, $g(x) < 0$ on (a,b), as required.

□

REMARK

(1) says that the graph of f lies above any tangent line, and (2) says that any secant line lies above the graph of f.

2-D Case

Suppose $f(x,y)$ has continuous second partial derivatives. We say that f is **strictly convex** if $Hf(x,y)$ is positive definite for all (x,y). By Theorem 2, f is strictly convex means $f_{xx} > 0$ and $f_{xx}f_{yy} - f_{xy}^2 > 0$ for all (x,y). We get a result which is analogous to Theorem 3.

THEOREM 4 If $f(x,y)$ has continuous second partial derivatives and is strictly convex, then

(1) $f(x,y) > L_{(a,b)}(x,y)$ for all $(x,y) \neq (a,b)$, and

(2) $f(a_1+t(b_1-a_1), a_2+t(b_2-a_2)) < f(a_1,a_2)+t[f(b_1,b_2)-f(a_1,a_2)]$ for $0 < t < 1$, $(a_1,a_2) \neq (b_1,b_2)$.

Proof: (1) Follows from Taylor's Theorem:

$$f = L_{(a,b)}(x,y) + \frac{1}{2}\left[f_{xx}(c,d)(x-a)^2 + 2f_{xy}(c,d)(x-a)(y-b) + f_{yy}(y-b)^2\right]$$

where (c,d) is on the line segment from (a,b) to (x,y). Since $f_{xx}(c,d) > 0$, $f_{xx}(c,d)f_{yy}(c,d) - f_{xy}(c,d)^2 > 0$, $R_{1,a,b}(x,y) > 0$ for $(x,y) \neq (a,b)$ by Theorem 2. Therefore, $f(x,y) > L_{(a,b)}(x,y)$ for $(x,y) \neq (a,b)$.

(2) We parameterize the line segment L from (a_1,a_2) to (b_1,b_2) by

$$L(t) = (a_1 + t(b_1 - a_1), a_2 + t(b_2 - a_2)), \qquad 0 \leq t \leq 1$$

For simplicity write $h = b_1 - a_1$ and $k = b_2 - a_2$. Define $g(t)$ by

$$g(t) = f(L(t)), \qquad 0 \le t \le 1 \tag{9.7}$$

Since f has continuous second partials by hypothesis, we can apply the Chain Rule to conclude that g' and g'' are continuous and are given by

$$g'(t) = f_x(L(t))h + f_y(L(t))k \tag{9.8}$$

$$g''(t) = f_{xx}(L(t))h^2 + 2f_{xy}(L(t))hk + f_{yy}(L(t))k^2 \tag{9.9}$$

for $0 \le t \le 1$. Since $f_{xx}(L(t)) > 0$ and $f_{xx}(L(t))f_{yy}(L(t)) - f_{xy}(L(t))^2 > 0$ for all t, $g''(t) > 0$ by Theorem 2. Thus, by Theorem 3, part (2):

$$g(t) < g(0) + \frac{g(1) - g(0)}{1 - 0}(t - 0), \qquad \text{for } 0 < t < 1$$

Therefore, $f(a_1 + t(b_1 - a_1), a_2 + t(b_2 - a_2)) < f(a_1, a_2) + t[f(b_1, b_2) - f(a_1, a_2)]$ for $0 < t < 1$ as required.

□

REMARK

(1) says that the graph of f lies above the tangent plane and (2) says that the cross-section of the graph of f above the line segment from (a_1, a_2) to (b_1, b_2) lies below the secant line.

EXAMPLE 5

If $f(x) = x^2$, then $f''(x) = 2 > 0$ for all x, so $f(x)$ is strictly convex.

If $f(x, y) = x^2 + y^2$, then $f_{xx} = 2$, $f_{xy} = 0$, and $f_{yx} = 2$, so $f_{xx}f_{yy} - f_{xy}^2 = 4 > 0$, so f is strictly convex.

THEOREM 5

If $f(x, y) \in C^2$ is strictly convex and has a critical point (c, d), then $f(x, y) > f(c, d)$ for all $(x, y) \neq (c, d)$ and f has no other critical point.

Proof: Note that $L_{(c,d)}(x, y) = f(c, d)$. Thus, $f(x, y) > f(c, d)$ for all $(x, y) \neq (c, d)$ by Theorem 4, part (1). If f has a second critical point (c_1, d_1), then by the reasoning just given $f(c_1, d_1) > f(c, d)$ and $f(c, d) > f(c_1, d_1)$ which is a contradiction. □

9.3 Proof of the Second Partial Derivative Test

We now want to prove part (1) of the second partial derivative test. The proof depends significantly on the hypothesis that the second partials of f are continuous, and on a plausible property of positive definite matrices: if you make a small change to the entries of a positive definite matrix then the new matrix is positive definite. This is proved separately as a lemma [1].

[1] This proof was provided by D. Siegel

LEMMA 1 Let $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ be a positive definite matrix. If $|\tilde{a} - a|$, $|\tilde{b} - b|$ and $|\tilde{c} - c|$ are sufficiently small, then $\begin{bmatrix} \tilde{a} & \tilde{b} \\ \tilde{b} & \tilde{c} \end{bmatrix}$ is positive definite.

Proof: Let Q and $\tilde{Q}$ be the quadratic forms determined by the given matrices i.e.

$$Q(u, v) = au^2 + 2buv + cv^2 \tag{9.10}$$

and similarly for $\tilde{Q}(u, v)$. We perform the change of variables

$$u = r\cos\theta, \qquad v = r\sin\theta$$

to obtain

$$Q(u, v) = r^2 p(\theta) \tag{9.11}$$

where

$$p(\theta) = a\cos^2\theta + 2b\cos\theta\sin\theta + c\sin^2\theta$$

Since for $r = 1$, $Q(u, v) = p(\theta)$, and Q is positive definite, we must have $p(\theta) > 0$ for all θ, $0 \le \theta \le 2\pi$.

Let

$$k = \min_{0\le\theta\le 2\pi} p(\theta)$$

Then $k > 0$ and by equation (9.11)

$$Q(u, v) \ge kr^2 \qquad \text{for all } (u, v) \ne (0, 0) \tag{9.12}$$

We are given that $|\tilde{a} - a|$, $|\tilde{b} - b|$ and $|\tilde{c} - c|$ are sufficiently small. Let

$$\delta = \max\{|\tilde{a} - a|, |\tilde{b} - b|, |\tilde{c} - c|\}$$

By equation (9.10) and the triangle inequality,

$$\begin{aligned} |Q(u, v) - \tilde{Q}(u, v)| &\le |\tilde{a} - a|u^2 + 2|\tilde{b} - b||u||v| + |\tilde{c} - c|v^2 \\ &\le \delta(u^2 + 2|u||v| + v^2) \\ &= \delta(|u| + |v|)^2 \\ &= \delta r^2(|\cos\theta| + |\sin\theta|)^2 \\ &< 4\delta r^2 \end{aligned}$$

We now choose $\delta = \frac{1}{8}k$. Then

$$|Q(u, v) - \tilde{Q}(u, v)| < \frac{1}{2}kr^2$$

which implies

$$\begin{aligned}\tilde{Q}(u,v) &\geq Q(u,v) - \frac{1}{2}kr^2 \\ &\geq kr^2 - \frac{1}{2}kr^2, \quad \text{by (9.12)} \\ &= \frac{1}{2}kr^2\end{aligned}$$

This shows that $\tilde{Q}(u,v) > 0$ for all $(u,v) \neq (0,0)$. Therefore, $\tilde{Q}(u,v)$ is positive definite. □

REMARK

The lemma is also true if "positive definite" is replaced by "negative definite" or "indefinite".

We now prove the second partial derivative test. For convenience we restate the theorem.

THEOREM 2 **(The Second Partial Derivative Test)**

Suppose that $f(x,y) \in C^2$ in some neighborhood of (a,b) and that

$$f_x(a,b) = 0 = f_y(a,b)$$

(1) If $Hf(a,b)$ is positive definite, then (a,b) is a local minimum point of f.

(2) If $Hf(a,b)$ is negative definite, then (a,b) is a local maximum point of f.

(3) If $Hf(a,b)$ is indefinite, then (a,b) is a saddle point of f.

Proof: We will prove (1).

We apply Taylor's formula with second order remainder. Since

$$f_x(a,b) = 0 = f_y(a,b)$$

Taylor's formula can be written as

$$f(x,y) - f(a,b) = \frac{1}{2}\Big[f_{xx}(c,d)(x-a)^2 + 2f_{xy}(c,d)(x-a)(y-b) + f_{yy}(c,d)(y-b)^2\Big] \quad (9.13)$$

where (c,d) lies on the line segment joining (a,b) and (x,y). The coefficient matrix in the quadratic expression on the right side of (9.13) is the Hessian matrix $Hf(c,d)$.

We are given that $Hf(a,b)$ is positive definite. By the lemma, there exists $\epsilon > 0$ such that if

$$|f_{xx}(x,y) - f_{xx}(a,b)| < \epsilon, |f_{xy}(x,y) - f_{xy}(a,b)| < \epsilon, |f_{yy}(x,y) - f_{yy}(a,b)| < \epsilon \quad (9.14)$$

then $Hf(x, y)$ is positive definite. Since the second partials of f are continuous at (a, b), the definition of continuity implies that there exists a $\delta > 0$ such that

$$\|(x, y) - (a, b)\| < \delta$$

implies (9.14) and hence that $Hf(x, y)$ is positive definite. Since

$$\|(c, d) - (a, b)\| < \|(x, y) - (a, b)\|$$

it follows that $Hf(c, d)$ is also positive definite. It now follows from equation (9.13) and the definition of positive definite matrix, that if $0 < \|(x, y) - (a, b)\| < \delta$, then $f(x, y) - f(a, b) > 0$. Thus, by definition (a, b) is a local minimum point of f. □

Parts (2) and (3) of the second derivative test can be proved in a similar way using the modified lemma.

Chapter 9 Problem Set

1. Find and classify the critical points of the function f, where
   (a) $f(x, y) = xy^2 - x^2y - xy + x^2$
   (b) $f(x, y) = xye^{x+2y}$
   (c) $f(x, y) = (x^2 + y^2 - 1)y$
   (d) $f(x, y) = x\sin(x + y)$
2. Find and classify the critical points of
$$f(x, y) = x^2 - 2x + y^3 - xy^2$$
3. Find and classify the critical points of
$$f(x, y) = (x + y)(xy + 1)$$
4. Find and classify the critical points of
$$f(x, y) = xy^2 + x^2y - 4xy$$
5. Find and classify the critical points of
$$f(x, y) = 4x^3 + 6x^2y + 3xy^2 - 3x$$
6. Find and classify the critical points of
$$f(x, y) = x^2 + y^2 + x^2y + 4$$
7. If the Hessian matrix is $\begin{pmatrix} 1 & -1 \\ -1 & 8 \end{pmatrix}$ at a critical point (a, b), then (a, b) is a local __________ of f (fill in the blank).
8. Find and classify all of the critical points for the following functions.
   (a) $f(x, y) = x^3 + y^3 - 3x^2 + 3y^2$
   (b) $f(x, y) = (x^2 + y^2)e^{-x^2}$
9. In each case invent a non-constant differentiable function $f(x, y)$ with the stated property. Classify the critical points of f, sketch the level curves, and describe the surface $z = f(x, y)$.
   (a) All points on the line $y = 2x$ are critical points of f.
   (b) All points on the circle $x^2 + y^2 = 1$ are critical points of f.
10. (a) Suppose that $f(x, y)$ is a C^2 function with one critical point (a, b) which has the property that $Hf(x, y)$ is positive definite for all (x, y), except possibly $(x, y) = (a, b)$. Prove that (a, b) is the unique *absolute minimum* for f on $\mathbb{R}^2$, i.e. $f(x, y) \geq f(a, b)$.
    (b) Let $f(x, y) = x^2 + y^2 + xy - x - 2y$. Show that f has one critical point and deduce that this is a unique global minimum.

Chapter 10

Optimization Problems

10.1 The Extreme Value Theorem

As we saw in Calculus 1, one is often interested in finding the largest or smallest possible value of a function f on some specified set S. We start with some standard definitions.

DEFINITION

Absolute Maximum and Minimum

Given a function $f(x, y)$ and a set $S \subseteq \mathbb{R}^2$,

1. a point $(a, b) \in S$ is an **absolute maximum point** of f on S if

$$f(x, y) \leq f(a, b) \quad \text{for all } (x, y) \in S$$

The value $f(a, b)$ is called the **absolute maximum value** of f on S.

2. a point $(a, b) \in S$ is an **absolute minimum point** of f on S if

$$f(x, y) \geq f(a, b) \quad \text{for all } (x, y) \in S$$

The value $f(a, b)$ is called the **absolute minimum value** of f on S.

The Extreme Value Theorem

Whether or not f has a maximum/minimum value on S depends on f and on the set S. Recall from Calculus 1 that the Extreme Value Theorem gives conditions which imply the existence of a maximum value and minimum value of f on an interval I. Here is the theorem.

THEOREM 1 **(The Extreme Value Theorem)**

If $f(x)$ is continuous on a finite closed interval I, then there exists $c_1, c_2 \in I$ such that

$$f(c_1) \leq f(x) \leq f(c_2) \quad \text{for all } x \in I$$

For our purposes, the important thing is to be able to give counterexamples to show that the conclusion may not be valid if the hypotheses are not satisfied.

EXERCISE 1

Give a function $f(x)$ and an interval I such that

1. I is closed, but f does not have an absolute maximum on I.
2. I is finite and f is continuous on I, but f does not have an absolute maximum on I.
3. I is infinite and f is continuous on I, but f does not have an absolute minimum.

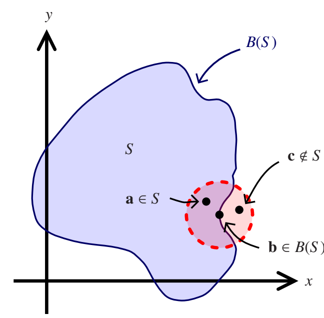
In order to generalize this theorem to functions of two variables, we need to generalize the concept of a finite closed interval to sets in $\mathbb{R}^2$.

DEFINITION
Bounded Set

A set $S \subset \mathbb{R}^2$ is said to be **bounded** if and only if it is contained in some neighbourhood of the origin.

Observe that the definition implies that every point in S must have finite distance from the origin.

Intuitively, a "boundary point" of a set $S \subset \mathbb{R}^2$ is a point which lies on the "edge" of S. Here is the definition.



DEFINITION
Boundary Point

Given a set $S \subset \mathbb{R}^2$, a point $(a, b) \in \mathbb{R}^2$ is said to be a **boundary point** of S if and only if every neighbourhood of (a, b) contains at least one point in S and one point not in S.

DEFINITION
Boundary of S

The set $B(S)$ of all boundary points of S is called the **boundary** of S.

DEFINITION
Closed Set

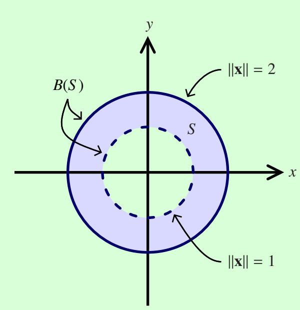
A set $S \subseteq \mathbb{R}^2$ is said to be **closed** if S contains all of its boundary points.

EXAMPLE 1

Consider $S = \{(x, y) \in \mathbb{R}^2 \mid 1 < \|(x, y)\| \leq 2\}$. The boundary of S is the set of all boundary points. So, as indicated in the diagram, the boundary of S is

$$B(S) = \{(x, y) \in \mathbb{R}^2 \mid \|(x, y)\| = 1 \text{ or } \|(x, y)\| = 2\}$$

Since the points (x, y) such that $\|(x, y)\| = 1$ are not in S, we have that S is not closed.



EXAMPLE 2 Consider $S = \{(x, y) \in \mathbb{R}^2 \mid x \geq 0\}$. The boundary of S is the y-axis which is in S. Therefore, S is closed.

Observe that the concept of a "closed set" in $\mathbb{R}^2$ generalizes the idea of a closed interval in $\mathbb{R}$.

We can now state the generalization of the Extreme Value Theorem to $\mathbb{R}^2$.

THEOREM 2 If $f(x, y)$ is continuous on a closed and bounded set $S \subset \mathbb{R}^2$, then there exists points $(a, b), (c, d) \in S$ such that

$$f(a, b) \leq f(x, y) \leq f(c, d) \quad \text{for all } (x, y) \in S$$

The proof is beyond the scope of this course.

Here are some counterexamples to show that the conclusion may not be valid if either hypothesis is not satisfied.

EXAMPLE 3 Let $S = \{(x, y) \mid x^2 + y^2 \leq 1\}$ and $f(x, y) = \begin{cases} 1 - x^2 - y^2 & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{if } (x, y) = (0, 0) \end{cases}$

Observe that S is the unit disc and hence is clearly bounded and it is closed since it contains its boundary, the circle $x^2 + y^2 = 1$. However, observe that for $(x, y) \in S$ we can make the values of f arbitrarily close to 1. But, since f is not continuous at $(0, 0)$, there is no $(x, y) \in S$ such that $f(x, y) = 1$. So, f does not have a maximum value on S.

EXAMPLE 4 Let $f(x, y) = x^2 + y^2$ and $S = \mathbb{R}^2$.

Clearly f is continuous on S. However, since S is not bounded, f does not have a maximum value on S. In particular, the values of f can be made arbitrarily large by increasing the values of x and/or y.

EXAMPLE 5 Let $f(x, y) = x^2 + y^2$ and $S = \{(x, y) \mid x^2 + y^2 < 1\}$

Clearly f is continuous on S. Observe that for $(x, y) \in S$ we can make the values of f arbitrarily close to 1. But, since S does not contain its boundary, there is no $(x, y) \in S$ such that $f(x, y) = 1$. Consequently, f does not have a maximum value on S.

REMARK

A function $f(x, y)$ may have an absolute maximum and/or an absolute minimum on a set $S \subseteq \mathbb{R}^2$ even if the conditions are not satisfied. We just cannot guarantee the existence with the theorem.

EXAMPLE 6 Let $S = \{(x, y) \in \mathbb{R}^2 \mid x > -1, y \in \mathbb{R}\}$ and let $f(x, y) = \begin{cases} 1 & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{if } (x, y) = (0, 0) \end{cases}$

Clearly f is not continuous on S and S is neither closed nor bounded. However, clearly 1 is the maximum of f on S and 0 is the minimum.

10.2 Algorithm for Extreme Values

Recall that if $f(x)$ is continuous, then the maximum value and minimum value of f on an interval $[a, b]$ occur either at a critical point of f (i.e. $f'(c) = 0$, or $f'(c)$ does not exist) or at an endpoint of the interval. Moreover, our algorithm for finding the maximum and/or minimum value for f on $[a, b]$ was to find the values of f at any critical points of f in $[a, b]$ and compare them to the values of f at the endpoints $x = a$ and $x = b$.

This approach can be generalized to $f(x, y)$. Let $S \subset \mathbb{R}^2$ be a closed and bounded set, with boundary $B(S)$ and suppose that f is continuous on S. The maximum value and minimum value of f on S occur either at a critical point of f that is in S, or at a point on the boundary of S. Thus, we get the following procedure which corresponds to what we were doing for functions of one variable in Calculus 1.

ALGORITHM

Let $S \subset \mathbb{R}^2$ be closed and bounded. To find the maximum and/or minimum value of a function $f(x, y)$ that is continuous on S,

(1) Find all critical points of f that are contained in S. Evaluate f at each such point.

(2) Find the maximum and minimum points of f on the boundary $B(S)$.

(3) The maximum value of f on S is the largest value of the function found in steps (1) and (2). The minimum value of f on S is the smallest value of the function found in steps (1) and (2).

REMARKS

1. The absolute maximum and/or minimum value may occur at more than one point in S.
2. It is not necessary to determine whether the critical points are local maximum or minimum points.

EXAMPLE 1 Find the maximum value of $f(x, y) = xy$ on the set

$$S = \{(x, y) \mid x^2 + y^2 \le 1\}$$

Solution: First, we observe that $\nabla f(x, y) = (y, x)$, hence the only critical point of f is $(0, 0)$ which is in S. We have $f(0, 0) = 0$.

Second, we look for the maximum value of f on the boundary $B(S)$ of S. To do this, we describe the boundary (the unit circle $x^2 + y^2 = 1$) in parametric form:

$$x = \cos t, \qquad y = \sin t, \qquad 0 \le t \le 2\pi$$

On $B(S)$, f has the values

$$g(t) = f(\cos t, \sin t) = \cos t \sin t = \frac{1}{2}\sin 2t$$

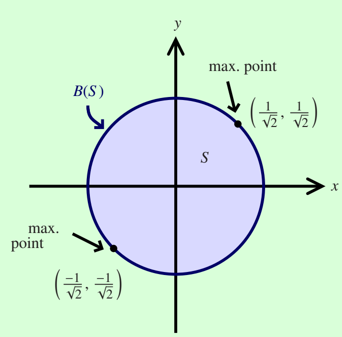
The problem now is to find the maximum value of $g(t)$ on the interval $0 \le t \le 2\pi$. We use the method from Calculus 1. We have

$$g'(t) = \cos 2t$$

Hence, on $0 \le t \le 2\pi$, the critical point of g are at $t = \frac{\pi}{4}, \frac{3\pi}{4}, \frac{5\pi}{4}, \frac{7\pi}{4}$. We have

$$g\left(\frac{\pi}{4}\right) = \frac{1}{2}, \quad g\left(\frac{3\pi}{4}\right) = -\frac{1}{2}, \quad g\left(\frac{5\pi}{4}\right) = \frac{1}{2}, \quad g\left(\frac{7\pi}{4}\right) = -\frac{1}{2}$$

Finally, we have $g(0) = 0$ and $g(2\pi) = 0$.



So, the maximum value of f on the boundary $B(S)$ is $\frac{1}{2}$ and occurs at $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$ and $\left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)$.

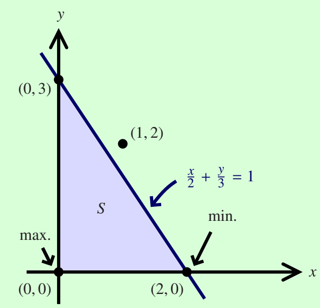
Comparing the values we found in the first and second step, we see that the maximum value of f on S is $\frac{1}{2}$ and occurs on the boundary at $\left(\pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}}\right)$.

EXERCISE 1 Find the maximum of $f(x, y) = x^2y - y$ on the set $S = \{(x, y) \mid 9x^2 + 4y^2 \leq 36\}$.

EXAMPLE 2 Find the maximum and minimum value of $f(x, y) = xy - 2x - y + 2$ on the triangular region S with vertices $(0, 0)$, $(2, 0)$ and $(0, 3)$.

Solution: First, we observe that $\nabla f(x, y) = (y - 2, x - 1)$ so the only critical point of f is $(1, 2)$. Since $(1, 2) \notin S$, this critical point plays no part in the solution.

The second step is to evaluate f on the boundary $B(S)$ of S. This has to be done on the three straight line segments separately. The values of f on $B(S)$ define a function of one variable, which we denote by g.



Case 1: $x = 0, 0 \leq y \leq 3$.

Let $g(y) = f(0, y) = -y + 2$. By inspection, the maximum of g on the interval $[0, 3]$ occurs at the end point $y = 0$, and the minimum occurs at the end point $y = 3$. So, $(0, 0)$ and $(0, 3)$ are possible maximum and minimum points for f.

Case 2: $y = 0, 0 \leq x \leq 2$.

Let $g(x) = f(x, 0) = -2x + 2$. As in Case 1, this leads to $(0, 0)$ and $(2, 0)$ as possible maximum and minimum points for f.

Case 3: $y = 3 - \frac{3}{2}x, 0 \leq x \leq 2$.

Let $g(x) = f(x, 3 - \frac{3}{2}x) = -\frac{3}{2}x^2 + \frac{5}{2}x - 1$, after simplifying. To find the critical points of g we solve

$$0 = g'(x) = -3x + \frac{5}{2}$$

This gives $x = \frac{5}{6}$. Hence, $\left(\frac{5}{6}, \frac{7}{4}\right)$ and the end points $(0, 3)$ and $(2, 0)$ are possible maximum and minimum points of f.

Now evaluate f at all points found above:

$$f(0, 0) = 2, \quad f(0, 3) = -1, \quad f(2, 0) = -2, \quad f\left(\frac{5}{6}, \frac{7}{4}\right) = \frac{1}{24}$$

Consequently, the maximum value of f on S is $f(0, 0) = 2$, and minimum value of f on S is $f(2, 0) = -2$.

EXERCISE 2 Find the maximum value of the function $f(x, y) = x^2y + xy^2$ on the triangular region with vertices $(0, 0)$, $(0, 1)$ and $(1, 0)$.

10.3 Optimization with Constraints

In many real world problems, we wish to find the maximum (minimum) of a function $f(x, y)$ subject to a constraint $g(x, y) = k$.

For example, assume that a manufacturer has three product lines. Let x, y, z denote the number of articles produced of each type and let a, b, c denote the profit per article for the three product lines respectively. The total profit is given by

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

Further assume that the manufacturer wishes to maintain production costs at a constant level k dollars per day. The production costs C depend on the number of articles x, y, z. That is, we require that

$$C(x, y, z) = k$$

The problem is to find the maximum profit $P(x, y, z)$ subject to the constraint $C(x, y, z) = k$.

Observe that in step 2 of our algorithm for finding extreme values, we also need to find the maximum and/or minimum of f subject to a constraint, namely the boundary of the region. In the last section, we did this by finding a parametric representation for the boundary. Of course, in many cases this may be extremely difficult or impossible to do.

So, we now derive an algorithm which will allow us to find the maximum and/or minimum of a differentiable function f on a smooth curve $g(x, y) = k$ without having to parameterize the curve.

Method of Lagrange Multipliers

We want to find the maximum (minimum) value of a differentiable function $f(x, y)$ subject to the constraint $g(x, y) = k$ where $g \in C^1$, or, more geometrically, find the maximum(minimum) value of $f(x, y)$ on the level set $g(x, y) = k$.

If $f(x, y)$ has a local maximum (or minimum) at (a, b) relative to nearby points on the curve $g(x, y) = k$ and $\nabla g(a, b) \neq (0, 0)$, then, by the Implicit Function Theorem (see Appendix A), $g(x, y) = k$ can be described by parametric equations

$$x = p(t), \quad y = q(t) \tag{10.1}$$

with p and q differentiable, and $(a, b) = (p(t_0), q(t_0))$ for some t_0. Define

$$u(t) = f(p(t), q(t))$$

The function u gives the values of f on the constraint curve, and hence has a local maximum (or minimum) at t_0. It follows that

$$u'(t_0) = 0 \tag{10.2}$$

Assuming f is differentiable we can apply the Chain Rule to get

$$u'(t) = f_x(p(t), q(t))p'(t) + f_y(p(t), q(t))q'(t)$$

Evaluating this at t_0 and using (10.2) gives

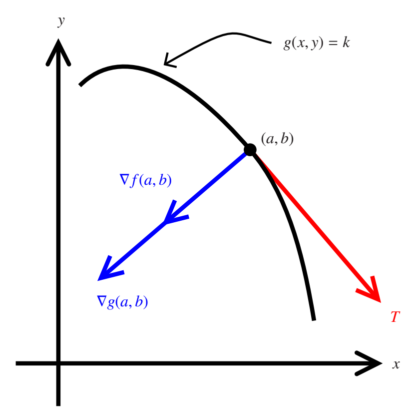
$$0 = f_x(a, b)p'(t_0) + f_y(a, b)q'(t_0)$$

This can be written as

$$\nabla f(a, b) \cdot (p'(t_0), q'(t_0)) = 0 \tag{10.3}$$

Recall the geometric interpretation of the gradient vector $\nabla g(a, b)$ proven in Theorem 7.2.2 that $\nabla g(a, b)$, if non-zero, is orthogonal to the level curve $g(x, y) = k$ at (a, b). Thus, since $(p'(t_0), q'(t_0))$ is the tangent vector to the constraint curve (10.1) we also have

$$\nabla g(a, b) \cdot \left(p'(t_0), q'(t_0)\right) = 0 \tag{10.4}$$



Since we are working in two dimensions, equations (10.3) and (10.4) imply that $\nabla f(a, b)$ and $\nabla g(a, b)$ are scalar multiples of each other. That is, there exists a constant λ such that

$$\nabla f(a, b) = \lambda \nabla g(a, b)$$

This leads to the following procedure, called the **Method of Lagrange Multipliers**.

ALGORITHM (Lagrange Multiplier Algorithm)

Assume that $f(x, y)$ is a differentiable function and $g \in C^1$. To find the maximum value and minimum value of f subject to the constraint $g(x, y) = k$, evaluate $f(x, y)$ at all points (a, b) which satisfy one of the following conditions.

(1) $\nabla f(a, b) = \lambda \nabla g(a, b)$ and $g(a, b) = k$.

(2) $\nabla g(a, b) = (0, 0)$ and $g(a, b) = k$.

(3) (a, b) is an end point of the curve $g(x, y) = k$.

The maximum/minimum value of $f(x, y)$ is the greatest/least value of f obtained at the points found in (1)-(3).

To find the points (a, b) in (1) we have to solve the system of 3 equations in 3 unknowns

$$\begin{aligned} f_x(x, y) &= \lambda g_x(x, y) \\ f_y(x, y) &= \lambda g_y(x, y) \\ g(x, y) &= k \end{aligned}$$

for x and y. It is important that this is done systematically so that you can ensure that you have found all possible points. We will demonstrate this in the examples below.

REMARKS

1. The variable λ, called the **Lagrange multiplier**, is not required for our purposes and so should be eliminated. However, in some real world applications, the value of λ can be extremely useful.

2. Case (2) and (3) are both exceptional. Observe that case (2) must be included since we assumed that $\nabla g(a, b) \neq (0, 0)$ in the derivation. Condition (3) will typically only arise if there are restrictions on the domain of $g(x, y)$ that result in the curve $g(x, y) = k$ having end points. For instance, if $g(x, y) = x^2 + y^2$, then the curve $g(x, y) = 1$ (the unit circle) does not have end points if there are no restrictions on x and y. However, if we restrict the domain of g to $\{(x, y) : y \geq 0\}$, then there will be two end points: $(-1, 0)$ and $(1, 0)$.

3. If the curve $g(x, y) = k$ is unbounded, then one must consider $\displaystyle\lim_{\|(x,y)\| \to \infty} f(x, y)$ for (x, y) satisfying $g(x, y) = k$.

EXAMPLE 1 Find the maximum value of $6x + 4y - 7$ on the ellipse $3x^2 + y^2 = 28$.

Solution: We want to find the maximum of

$$f(x, y) = 6x + 4y - 7$$

subject to the constraint

$$g(x, y) = 3x^2 + y^2 = 28$$

(1) $\nabla f(x, y) = \lambda \nabla g(x, y), \quad g(x, y) = 28.$

Differentiating gives $\nabla f(x, y) = (6, 4)$ and $\nabla g(x, y) = (6x, 2y)$. Comparing entries of $\nabla f(x, y) = \lambda \nabla g(x, y)$ and adding the constraint equation $g(x, y) = 28$ gives the system of equations

$$6 = 6\lambda x \tag{10.5}$$
$$4 = 2\lambda y \tag{10.6}$$
$$3x^2 + y^2 = 28 \tag{10.7}$$

By equation (10.5), $x \neq 0$ and so $\lambda = \frac{1}{x}$, which when substituted in equation (10.6) gives $y = 2x$. We substitute this into (10.7) and solve for x, obtaining $x = \pm 2$. For $x = 2$ we get $y = 2(2) = 4$, for $x = -2$ we get $y = 2(-2) = -4$. Thus, we obtain two points $(2, 4)$ and $(-2, -4)$.

(2) $\nabla g(x, y) = (0, 0), \quad g(x, y) = 28.$

We have $(0, 0) = \nabla g(x, y) = (6x, 2y)$ implies $x = y = 0$, which does not satisfy the constraint (10.7). Hence, there are no points in this step.

(3) Check end points.

There are no endpoints since the constraint is a closed curve (an ellipse).

Finally, we evaluate f at all the points found in the above 3 steps. We get

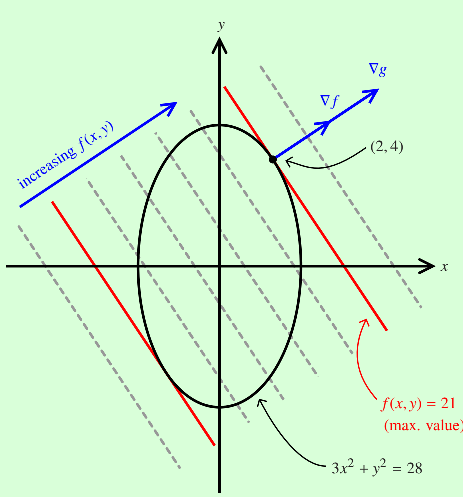
$$f(2, 4) = 21$$
$$f(-2, -4) = -35$$

So, the maximum value of f on $3x^2 + y^2 = 28$ is 21 and occurs at $(2, 4)$.

We can view the result geometrically. The straight lines are the level curves

$$f(x, y) = 6x + 4y - 7 = k$$

Notice that ∇f and ∇g are parallel at the maximum point.



EXAMPLE 2 Find the maximum and minimum values of $f(x, y) = y$ on the piriform curve defined by $y^2 + x^4 - x^3 = 0$.

Solution: We have $f(x, y) = y$ and constraint $g(x, y) = y^2 + x^4 - x^3 = 0$.

(1) $\nabla f(x, y) = \lambda \nabla g(x, y), \quad g(x, y) = 0$.

We get $\nabla f(x, y) = (0, 1)$ and $\nabla g(x, y) = (4x^3 - 3x^2, 2y)$, so we need to solve

$$0 = \lambda(4x^3 - 3x^2) = x^2(4x - 3)\lambda \tag{10.8}$$
$$1 = \lambda(2y) \tag{10.9}$$
$$0 = y^2 + x^4 - x^3 \tag{10.10}$$

Clearly $\lambda \neq 0$ because of (10.9), so (10.8) gives $x = 0$ or $x = \frac{3}{4}$.

If $x = 0$, then (10.10) gives $y = 0$ which does not satisfy (10.9).

If $x = \frac{3}{4}$, then (10.10) gives $0 = y^2 - \frac{27}{256}$ which implies $y = \pm\frac{3\sqrt{3}}{16}$. Hence, we get two points $\left(\frac{3}{4}, \frac{3\sqrt{3}}{16}\right)$ and $\left(\frac{3}{4}, -\frac{3\sqrt{3}}{16}\right)$.

(2) $\nabla g(x, y) = (0, 0), \quad g(x, y) = 0$.

We have

$$(0, 0) = \nabla g(x, y) = (4x^3 - 3x^2, 2y)$$

which implies $0 = 4x^3 - 3x^2 = x^2(4x - 3)$ and $2y = 0$. So, we get points $(0, 0)$, and $\left(\frac{3}{4}, 0\right)$. However, $\left(\frac{3}{4}, 0\right)$ is not on the constraint curve, so we just have one point $(0, 0)$.

(3) Check end points.

Graphing the piriform curve, we see that it is closed. So, there are no end points.

Evaluating f at all the points found above gives

$$f\left(\frac{3}{4}, -\frac{3\sqrt{3}}{16}\right) = -\frac{3\sqrt{3}}{16}$$
$$f\left(\frac{3}{4}, \frac{3\sqrt{3}}{16}\right) = \frac{3\sqrt{3}}{16}$$
$$f(0, 0) = 0$$

Thus, the maximum value is $\frac{3\sqrt{3}}{16}$ at $\left(\frac{3}{4}, \frac{3\sqrt{3}}{16}\right)$ and the minimum value is $-\frac{3\sqrt{3}}{16}$ at $\left(\frac{3}{4}, -\frac{3\sqrt{3}}{16}\right)$.

EXAMPLE 3 Let R be the region bounded by the curve $x = \sqrt{1-y^2}$ and the y-axis. Find the maximum and minimum value of $f(x, y) = x^2 - \frac{1}{2}x + y^2$ on the region R.

Solution: Observe that this is an extreme values on a *region* problem as in Section 10.2. Thus, we apply our algorithm from Section 10.2.

We first find critical points of f inside the region R. We have

$$\nabla f = (2x - \frac{1}{2}, 2y) = (0, 0) \Rightarrow x = \frac{1}{4}, y = 0$$

There is one critical point $\left(\frac{1}{4}, 0\right)$, which is inside the region and $f\left(\frac{1}{4}, 0\right) = -\frac{1}{16}$.

Next, we find the maximum and minimum of f on the boundary of R. The boundary has two parts, the y-axis and the semi-circle $x = \sqrt{1-y^2}$.

For the y-axis, we have $x = 0$, $-1 \leq y \leq 1$, so on this line we have $f(0, y) = 0 + y^2$ which we know has minimum 0 at $(0, 0)$ and maximum 1 and $(0, \pm 1)$.

For the semi-circle, instead of parameterizing it as we did in Section 10.2, we will use the method of Lagrange multipliers. To make the calculations easier, we simplify the constraint to $x^2 + y^2 = 1$, $x \geq 0$. Hence, we take $g(x, y) = x^2 + y^2 = 1$, $x \geq 0$.

(1) $\nabla f(x, y) = \lambda \nabla g(x, y)$, $g(x, y) = 1$.

$$2x - \frac{1}{2} = \lambda(2x) \tag{10.11}$$

$$2y = \lambda(2y) \tag{10.12}$$

$$x^2 + y^2 = 1, \qquad x \geq 0 \tag{10.13}$$

From (10.12) we see that $y = 0$ or $\lambda = 1$.

If $y = 0$, then (10.13) gives $x = 1$ (since $x \geq 0$). With $\lambda = \frac{3}{4}$ (10.11) is also satisfied. Thus, $(1, 0)$ is a point.

If $\lambda = 1$, then (10.11) is $2x - \frac{1}{2} = 2x$, which has no solutions, so we get no points.

(2) $\nabla g(x, y) = (0, 0)$, $g(x, y) = 1$.

We have $\nabla g(x, y) = (2x, 2y) = (0, 0)$ only if $x = 0$ and $y = 0$, but this does not satisfy the constraint so no points.

(3) Check end points.

The semi-circle has end points when $x = 0$, so at $(0, 1)$ and $(0, -1)$.

Putting all the points into f gives

$$f(1, 0) = \frac{1}{2}, \quad f(0, 1) = 1, \quad f(0, -1) = 1$$

Thus, on the semi-circle the maximum of f is 1 at $(0, \pm 1)$ and the minimum of f is $\frac{1}{2}$ at $(1, 0)$.

Comparing the values of f found in all steps, we find that the maximum of f on R is 1 at $(0, \pm 1)$ and the minimum of f is $-\frac{1}{16}$ at $\left(\frac{1}{4}, 0\right)$.

EXERCISE 1 Find the maximum value of xy on the circle $x^2 + y^2 = 1$. Sketch the constraint curve and some level curves of xy showing the gradient vectors at the maximum point.

EXERCISE 2 Find the maximum and minimum value of $F(x, y) = x^2 + 2x + y^2$ subject to the constraint $x^2 + 4y^2 \leq 24$.

Functions of Three Variables

We can generalize the algorithm for $f(x, y)$ to work for functions of three variables $f(x, y, z)$.

ALGORITHM

To find the maximum/minimum value of a differentiable function $f(x, y, z)$ subject to $g(x, y, z) = k$ such that $g \in C^1$, we evaluate $f(x, y, z)$ at all points (a, b, c) which satisfy one of the following:

(1) $\nabla f(a, b, c) = \lambda \nabla g(a, b, c)$ and $g(a, b, c) = k$.

(2) $\nabla g(a, b, c) = (0, 0, 0)$ and $g(a, b, c) = k$.

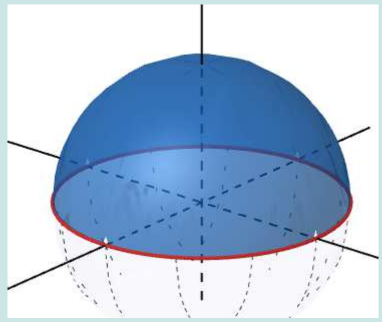
(3) (a, b, c) is an edge point of the surface $g(x, y, z) = k$. (See Remark below.)

The maximum/minimum value of $f(x, y, z)$ is the largest/smallest value of f obtained from all points found in (1)-(3).

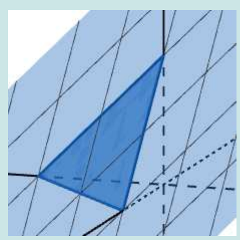
REMARK

1. If condition (1) in the algorithm holds, it follows that the level surface $f(x, y, z) = f(a, b, c)$ and the constraint surface $g(x, y, z) = k$ are tangent at the point (a, b, c), since their normals coincide (see Theorem 7.3.1).

2. The surface $g(x, y, z) = k$ will usually not have any edge points. Edge points typically arise when we introduce restrictions to the domain of g. We will not formally define what an edge point is; instead, we shall illustrate with examples. Let $g(x, y, z) = x^2+y^2+z^2$. Then the surface $g(x, y, z) = 1$ (the unit sphere) has no edge points. However, if we restrict the domain of g to $\{(x, y, z) : z \geq 0\}$, then the resulting surface (the upper unit hemisphere) will have edge points all along the unit circle in the xy-plane:



Here is another example. If $g(x, y, z) = x + y + z$ then the surface $g(x, y, z) = 3$ is a plane in $\mathbb{R}^3$. It has no edge points. If we restrict the domain of g to $\{(x, y, z) : x, y, z \geq 0\}$, then the resulting surface will be the portion of the plane in the first octant. This is a triangle. It has edge points along each of its three boundary edges:



EXAMPLE 4 Find the point on the sphere $x^2 + y^2 + z^2 = 1$ which is closest to the point $(1, 2, 2)$.

Solution: We want to minimize the distance between the point $(1, 2, 2)$ and a point (x, y, z) on the given sphere. To simplify things, we consider the square of this distance, which is given by the function

$$f(x, y, z) = (x - 1)^2 + (y - 2)^2 + (z - 2)^2$$

The constraint is $g(x, y, z) = x^2 + y^2 + z^2 = 1$.

(1) $\nabla f(x, y, z) = \lambda \nabla g(x, y, z)$, $g(x, y, z) = 1$.

$$2(x-1) = 2\lambda x \tag{10.14}$$
$$2(y-2) = 2\lambda y \tag{10.15}$$
$$2(z-2) = 2\lambda z \tag{10.16}$$
$$x^2 + y^2 + z^2 = 1 \tag{10.17}$$

Observe that (10.14), (10.15), and (10.16) give that $x \neq 0$, $y \neq 0$, and $z \neq 0$. Hence, solving these equations for λ and setting them equal to each other gives

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

Looking at each pair, we find that $y = 2x$, $z = 2x$, and thus $y = z$. Putting these into the constraint (10.17) gives two points, $\left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right)$ and $\left(-\frac{1}{3}, -\frac{2}{3}, -\frac{2}{3}\right)$.

(2) $\nabla g(x, y, z) = (0, 0, 0)$, $g(x, y, z) = 1$.

We have $\nabla g(x, y, z) = (0, 0, 0)$ implies $x = y = z = 0$, which does not satisfy the constraint.

(3) Edge points

There are no edge points on the unit sphere.

Evaluating f at all the points found above gives

$$f\left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right) = 4$$
$$f\left(-\frac{1}{3}, -\frac{2}{3}, -\frac{2}{3}\right) = 16$$

Thus, the point $\left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right)$ is the point on the sphere $x^2 + y^2 + z^2 = 1$ that is closest to the point $(1, 2, 2)$.

REMARK

Keep in mind the geometric interpretation. The level sets $f(x, y, z) = k$ are spheres centered on the point $(1, 2, 2)$. The minimum distance occurs when one of the spheres touches (i.e. is tangent to) the constraint surface which is the sphere $g(x, y, z) = 1$. At the point of tangency the normals are parallel, i.e. $\nabla f = \lambda \nabla g$.

EXERCISE 3 Find the points on the surface $z^2 = xy + 1$ that are closest to the origin.

Generalization

The method of Lagrange multipliers can be generalized to functions of n variables $f(\mathbf{x})$, $\mathbf{x} \in \mathbb{R}^n$ and with r constraints of the form

$$g_1(\mathbf{x}) = 0, \quad g_2(\mathbf{x}) = 0, \quad \ldots, \quad g_r(\mathbf{x}) = 0 \tag{10.18}$$

In order to find the possible maximum and minimum points of f subject to the constraints (10.18), one has to find all points $\mathbf{a}$ such that

$$\nabla f(\mathbf{a}) = \lambda_1 \nabla g_1(\mathbf{a}) + \cdots + \lambda_r \nabla g_r(\mathbf{a}), \quad \text{and} \quad g_i(\mathbf{a}) = 0, \qquad 1 \le i \le r$$

The scalars $\lambda_1, \ldots, \lambda_r$ are the Lagrange multipliers. When $r = 1$, and $n = 2$ or 3, this reduces to the previous algorithms.

Chapter 10 Problem Set

1. Find the maximum and minimum values of the function $f(x, y) = xy - x^3y^2$ on the square $0 \le x \le 1$, $0 \le y \le 1$.
2. Find the maximum and minimum values of the function $f(x, y) = x + 2y$ on the disc $x^2 + y^2 \le 4$.
3. Find the maximum and minimum values of the function $f(x, y) = xye^{-\frac{1}{2}x-\frac{1}{3}y}$ on the triangular set with vertices $(0, 0)$, $(2, 0)$ and $(0, 3)$.
4. Find the maximum and minimum of the function $f(x, y) = x^3 - 3x + y^2 + 2y$ on the region bounded by the lines $x = 0$, $y = 0$, $x + y = 1$.
5. The steady-state temperature at position (x, y) of a metal disc, $x^2 + y^2 \le b^2$, where b is a positive constant, is given by

   $$f(x, y) = 100 + x^3 - 3xy^2$$

   Find the hottest and coldest points on the disc.
6. (a) Use Lagrange multipliers to find the greatest and least distance of the curve $6x^2 + 4xy + 3y^2 = 14$ from the origin.

   (b) Illustrate the result graphically by drawing the constraint curve $g(x, y) = 0$, the level curves $f(x, y) = C$, and the gradient vectors ∇f and ∇g. Clearly indicate the relation between the level curves of f and the constraint curve at the maximum and minimum.
7. Assume the earth is located at $(x, y, z) = (0, 0, 0)$ and the path of a comet is given by

   $$3x^2 + 8xy - 3y^2 = 5^3, \quad z = 0, \quad x > 0$$

   Find the distance of closest approach to the centre of the earth. Units are in $km \times 10^5$. Illustrate your answer with a sketch.
   Suggestion: In order to avoid messy square roots, use the method of Lagrange multipliers.
8. Use the method of Lagrange multipliers to find the maximum and minimum values of $xy + z^2$ on the surface $x^2 + y^2 + z^2 = 1$.
9. Use Lagrange multipliers to find the maximum value of $x + y + z$ on the ellipsoid $x^2 + \frac{1}{4}y^2 + \frac{1}{9}z^2 = 1$. Discuss briefly a geometrical interpretation.
10. Solve questions 2 and 5, using Lagrange multipliers as part of your solution.
11. Let $f(x, y) = x^2 + y^2 - \frac{1}{2}y$.

    (a) Use the method of Lagrange multipliers to find the maximum and minimum points of $f(x, y)$ on the curve $y = \sqrt{1 - 2x^2}$.

    (b) Let R be the region bounded by the curve $y = \sqrt{1 - 2x^2}$ and the x-axis. Find the maximum and minimum value of $f(x, y)$ on the region R.
12. Find the greatest and least distance of the surface $6x^2 + 4xy + 3y^2 + 14z^2 = 14$ from the origin.
13. Use the method of Lagrange multipliers to find the maximum and minimum values of $f(x, y) = x$ on the piriform curve defined by

    $$y^2 + x^4 - x^3 = 0$$
14. An open irrigation channel is to be made in symmetric form with 3 straight sides, as drawn.

If the sum of the lengths of the sides of the cross-section equals L (given), find the channel design which will permit the maximum possible flow.

Comment: You should formulate the problem mathematically in the form: find the maximum value of a function on a closed and bounded subset of $\mathbb{R}^2$.

15. Consider all pentagons which have a line of symmetry, two adjacent interior angles of 90^0, and a perimeter of fixed length L. Find the shape that encloses the largest area.

16. Find the maximum and minimum value of the function $f(x, y) = (x + 1)^2 + y^2$ on the part of the graph of $y^2 - x^3 = 0$ from $(1, -1)$ to $(1, 1)$.

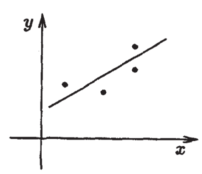
17. * Prove that $x^4 + y^4 - 4b^2xy \geq -2b^4$ for all $x, y \in \mathbb{R}$.

18. Consider $f(x, y) = (x^2 + y^2 + k)e^{-x^2-y^2}$ where k is a constant. The properties of f depend in a significant way on k. Analyse the function as regards local and global maxima and minima. Sketch/describe the surface $z = f(x, y)$. How many qualitatively different cases are there?

19. Consider a set of points (x_i, y_i), $i = 1, 2, \ldots, n$, which are close to lying on a straight line $y = mx + b$. In order to find the straight line which "best fits" the points, we minimize the sum of the squares of the errors:

$$E(m, b) = \sum_{i=1}^{n} [y_i - (mx_i + b)]^2$$

In other words, we find the minimum value of $E(m, b)$, for all values of the slope m and intercept b, i.e. for all $(m, b) \in \mathbb{R}^2$.



Apply this method to find the straight line which best fits the points $(0, 1)$, $(2, 3)$, $(3, 6)$ and $(4, 8)$.

Suggestion: Do not expand $E(m, b)$ before calculating the partial derivatives.

20. * Suppose that a function $f(x, y)$ has exactly one critical point which is a local minimum. Does f have a minimum on $\mathbb{R}^2$? Discuss with reference to the functions $f_1(x, y) = x^2 + y^2(1 - x)^3$ and $f_2(x, y) = x^2 + y^2$.

21. * (a) Use the method of Lagrange multipliers to prove that if $x_1^2 + x_2^2 + x_3^2 = 1$, then

$$x_1^2 x_2^2 x_3^2 \leq \frac{1}{3^3}$$

(b) Hence prove that for all positive real numbers a_1, a_2 and a_3,

$$(a_1 a_2 a_3)^{\frac{1}{3}} \leq \frac{a_1 + a_2 + a_3}{3}$$

(c) Generalize (a) and (b) to deduce the *arithmetic-geometric mean inequality*:

$$(a_1 a_2 \cdots a_n)^{\frac{1}{n}} \leq \frac{a_1 + a_2 + \cdots + a_n}{n}$$

for all positive real numbers $a_1, a_2, \ldots, a_n$ and any positive integer n.

Chapter 11

Double Integrals

11.1 Definition of the Double Integral

Recall that to find the area under a continuous curve $y = f(x)$ over a closed interval $[a, b]$ we used a single integral which we defined as a limit of Riemann sums:

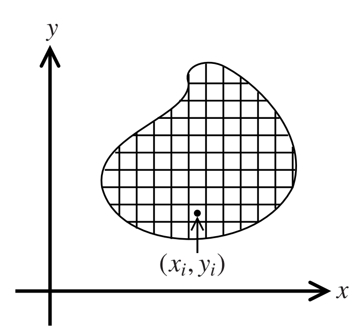
$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \sum_{i=1}^{n} f(x_i)\Delta x_i$$

where Δx_i is the length of the i-th subinterval in some decomposition (i.e. partition) of the interval $[a, b]$ and x_i is some point in the i-th subinterval.

We found that the single integral had many applications beside calculating areas under curves. We can use single integrals for finding mass of thin rods, calculating work, and for finding volumes of revolution. However, what if we want to calculate the mass of a thin plate, or to find the volume of more complicated regions? For these, we use double integrals.

Let D be a closed and bounded set in $\mathbb{R}^2$ whose boundary is a piecewise smooth closed curve. Let $f(x, y)$ be a function which is bounded on D, that is, there exists a number M such that $|f(x, y)| \leq M$ for all $(x, y) \in D$.

Subdivide D by means of straight lines parallel to the axes, forming a partition P of D. Label the n rectangles that lie completely in D, in some specific order, and denote their areas by ΔA_i, $i = 1, \ldots, n$. Choose a point (x_i, y_i) in the i-th rectangle and form the Riemann sum



$$\sum_{i=1}^{n} f(x_i, y_i)\Delta A_i \qquad (11.1)$$

DEFINITION
Integrable

Let $D \subset \mathbb{R}^2$ be closed and bounded. Let P be a partition of D as described above, and let $|\Delta P|$ denote the length of the longest side of all rectangles in the partition P. A function $f(x, y)$ which is bounded on D is **integrable** on D if all Riemann sums approach the same value as $|\Delta P| \to 0$.

DEFINITION
Double Integral

If $f(x, y)$ is integrable on a closed bounded set D, then we define the **double integral** of f on D as

$$\iint\limits_D f(x, y)\, dA = \lim_{\Delta P \to 0} \sum_{i=1}^{n} f(x_i, y_i)\Delta A_i$$

Is there any guarantee that the limiting process in the definition of the double integral actually leads to a unique value, i.e. that the limit exists? It is possible to define weird functions for which the limit does not exist, i.e. which are not integrable on D. However, if f is continuous on D, it can be proved that f is integrable on D, that is the double integral of f does exist. Functions which are discontinuous on D may be integrable on D. For example, if f is continuous in D except at points which lie on a curve C (f is piece-wise continuous), then f is integrable. The proofs of these results are beyond the scope of this course.

Interpretation of the Double Integral

When you encounter the double integral symbol

$$\iint\limits_D f(x, y)\, dA$$

think of "limit of a sum". In itself, the double integral is a mathematically defined object. It has many interpretations depending on the meaning that you assign to the integrand $f(x, y)$. The "dA" in the double integral symbol should remind you of the area of a rectangle in a partition of D.

Double Integral as Area:

The simplest interpretation is when you specialize f to be the constant function with value unity:

$$f(x, y) = 1, \quad \text{for all } (x, y) \in D$$

Then the Riemann sum (14.1) simply sums the areas of all rectangles in D, and the double integral serves to define the area $A(D)$ of the set D:

$$A(D) = \iint\limits_D 1\, dA$$

Double Integral as Volume:

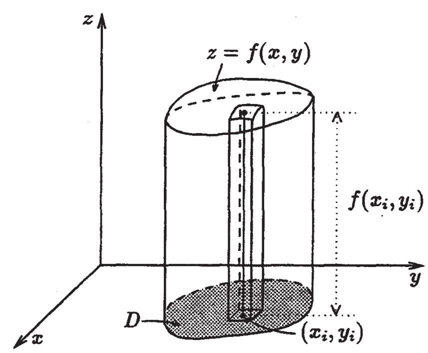
If $f(x, y) \geq 0$ for all $(x, y) \in D$, then the double integral

$$\iint_D f(x, y)\, dA$$

can be interpreted as the volume $V(S)$ of the region defined by

$$S = \left\{(x, y, z) \mid 0 \leq z \leq f(x, y), (x, y) \in D\right\}$$

which represents the solid below the surface $z = f(x, y)$ and above the set D in the xy-plane. The justification is as follows.



The partition P of D decomposes the solid S into vertical "columns". The height of the column above the i-th rectangle is approximately $f(x_i, y_i)$, and so its volume is approximately

$$f(x_i, y_i)\Delta A_i$$

The Riemann sum (14.1) thus approximates the volume $V(S)$:

$$V(S) \approx \sum_{i=1}^{n} f(x_i, y_i)\Delta A_i$$

As $|\Delta P| \to 0$ the partition becomes increasingly fine, so the error in the approximation will tend to zero. Thus, the volume $V(S)$ is

$$V(S) = \iint_D f(x, y)\, dA$$

Double Integral as Mass:

Think of a thin flat plate of metal whose density varies with position. Since the plate is thin, it is reasonable to describe the varying density by an "area density", that is a function $f(x, y)$ that gives the mass per unit area at position (x, y). In other words, the mass of a small rectangle of area ΔA_i located at position (x_i, y_i) will be approximately

$$\Delta M_i \approx f(x_i, y_i)\Delta A_i$$

The Riemann sum (14.1) corresponding to a partition P of D will approximate the total mass M of the plate D, and the double integral of f over D, being the limit of the sum, will represent the total mass:

$$M = \iint_D f(x, y)\, dA$$

Double Integral as Probability:

Let $f(x, y)$ be the probability density of a continuous 2-D random variable (X, Y). The probability that $(X, Y) \in D$, a given subset of $\mathbb{R}^2$, is

$$P((X, Y) \in D) = \iint\limits_D f(x, y)\, dA$$

Average Value of a Function:

The double integral is also used to define the average value of a function $f(x, y)$ over a set $D \subset \mathbb{R}^2$.

Recall for a function of one variable, $f(x)$, the average value of f over an interval $[a, b]$, denoted f_{av}, is defined by

$$f_{av} = \tfrac{1}{b-a} \int_a^b f(x)\, dx$$

Similarly, for a function of two variables $f(x, y)$, we can define the average value of f over a closed and bounded subset D of $\mathbb{R}^2$ by

$$f_{av} = \tfrac{1}{A(D)} \iint\limits_D f(x, y)\, dA$$

EXERCISE 1

A city occupies a region D of the xy-plane. The population density in the city (measured as people/unit area) depends on position (x, y), and is given by a function $p(x, y)$. Interpret the double integral $\iint\limits_D p(x, y)\, dA$.

Properties of the Double Integral

The basic properties of single integrals can be generalized to double integrals. We do not give the proofs, but we point out that the results are plausible if one thinks in terms of Riemann sums.

THEOREM 1

(Linearity)

If $D \subset \mathbb{R}^2$ is a closed and bounded set and f and g are two integrable functions on D, then for any constant c:

$$\iint\limits_D (f + g)\, dA = \iint\limits_D f\, dA + \iint\limits_D g\, dA$$

$$\iint\limits_D cf\, dA = c \iint_D f\, dA$$

THEOREM 2 **(Basic Inequality)**

If $D \subset \mathbb{R}^2$ is a closed and bounded set and f and g are two integrable functions on D such that $f(x, y) \le g(x, y)$ for all $(x, y) \in D$, then

$$\iint_D f\, dA \le \iint_D g\, dA$$

THEOREM 3 **(Absolute Value Inequality)**

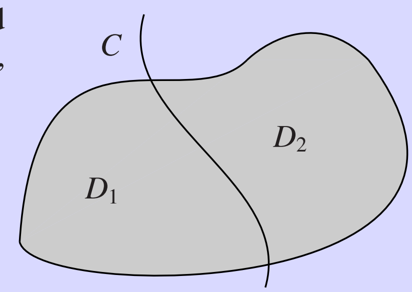
If $D \subset \mathbb{R}^2$ is a closed and bounded set and f is an integrable function on D, then

$$\left| \iint_D f\, dA \right| \le \iint_D |f|\, dA$$

THEOREM 4 **(Decomposition)**

Assume $D \subset \mathbb{R}^2$ is a closed and bounded set and f is an integrable function on D. If D is decomposed into two closed and bounded subsets D_1 and D_2 by a piecewise smooth curve C, then

$$\iint_D f\, dA = \iint_{D_1} f\, dA + \iint_{D_2} f\, dA$$



$D = D_1 \cup D_2$

REMARKS

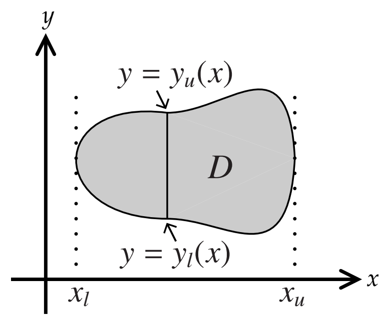
1. The Basic Inequality can be used to obtain an estimate for a double integral that cannot be evaluated exactly.
2. The decomposition property is essential for dealing with complicated regions of integration and with discontinuous integrands.

11.2 Iterated Integrals

It is clear that double integrals can be evaluated approximately by using a computer to evaluate a suitable Riemann sum. The accuracy would depend on how fine a partition you choose. But it is natural to ask: is it possible to calculate double integrals exactly, using methods that work for single integrals? For sufficiently simple functions and regions of integration, the answer is YES. The idea is to write the double integral as a succession of two single integrals, called an **iterated integral**. We will derive a method for doing this by using the interpretation of the double integral as volume.

Let D be a region in the xy-plane and let f be a function such that $f(x, y) \geq 0$ for all $(x, y) \in D$. If V denotes the volume of the solid above D and below the surface $z = f(x, y)$, then we have

$$V = \iint_D f(x, y)\, dA$$



Assume that the region D lies between vertical lines $x = x_\ell$ and $x = x_u$ with $x_\ell < x_u$ and has top curve $y = y_u(x)$ and bottom curve $y = y_\ell(x)$. That is, D is described by the inequalities

$$y_\ell(x) \leq y \leq y_u(x), \quad \text{and} \quad x_\ell \leq x \leq x_u$$

Now, recall from Calculus 2 that we can find a volume of a region by integrating over all possible cross-sectional areas. That is,

$$V = \int_{x_\ell}^{x_u} A(x)\, dx$$

where $A(x)$ is the cross-sectional area of the solid for any fixed value of x. But, we know that the cross-sectional area $A(x)$ is the area under the cross-section $z = f(x, y)$, and thus is given by a single integral

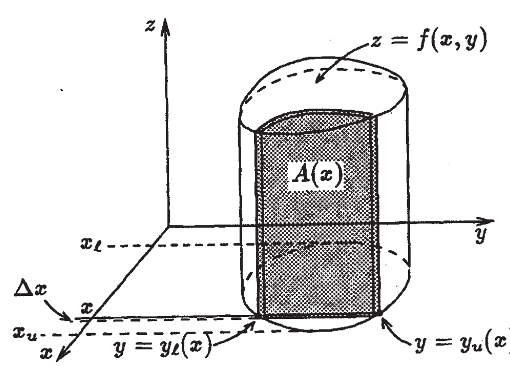
$$A(x) = \int_{y_\ell(x)}^{y_u(x)} f(x, y)\, dy$$

Hence, the volume of the region is

$$V = \int_{x_\ell}^{x_u} \left(\int_{y_\ell(x)}^{y_u(x)} f(x, y)\, dy \right) dx$$

Thus, we have

$$\iint_D f(x, y)\, dA = \int_{x_\ell}^{x_u} \int_{y_\ell(x)}^{y_u(x)} f(x, y)\, dy\, dx$$



as desired.

THEOREM 1

Let $D \subset \mathbb{R}^2$ be defined by

$$y_\ell(x) \le y \le y_u(x), \quad \text{and} \quad x_\ell \le x \le x_u$$

where $y_\ell(x)$ and $y_u(x)$ are continuous for $x_\ell \le x \le x_u$. If $f(x, y)$ is continuous on D, then

$$\iint\limits_D f(x, y)\, dA = \int_{x_\ell}^{x_u} \int_{y_\ell(x)}^{y_u(x)} f(x, y)\, dy\, dx$$

The proof is beyond the scope of this course.

REMARK

Although the parentheses around the inner integral are usually omitted, we must evaluate it first. Moreover, as in our interpretation of volume above, when evaluating the inner integral, we are integrating with respect to y while holding x constant. That is, we are using **partial integration**.

EXAMPLE 1

Evaluate

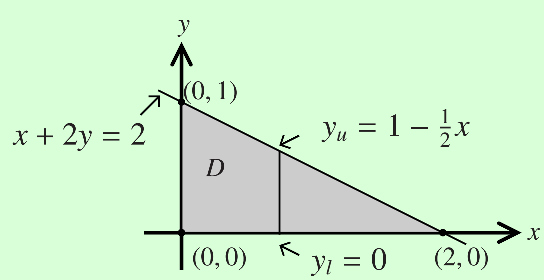
$$\iint\limits_D xy\, dA$$

where D is the triangular region with vertices $(0, 0)$, $(2, 0)$, and $(0, 1)$.

Solution: The set D is defined by

$$0 \le y \le 1 - \frac{1}{2}x, \quad \text{and } 0 \le x \le 2$$

$$\begin{aligned}
\iint\limits_D xy\, dA &= \int_{x=0}^{2} \int_{y=0}^{1-\frac{1}{2}x} xy\, dy\, dx \\
&= \int_{x=0}^{2} x\left(\frac{1}{2}y^2\right)\Big|_0^{1-\frac{1}{2}x} dx \\
&= \frac{1}{2}\int_0^2 x\left(1 - \frac{1}{2}x\right)^2 dx \\
&= \frac{1}{4}x^2 - \frac{1}{6}x^3 + \frac{1}{32}x^4\Big|_0^2 \\
&= \frac{1}{6}
\end{aligned}$$

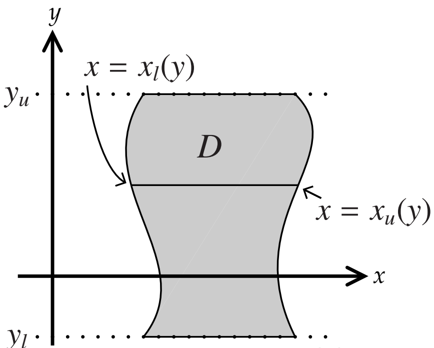


Suppose now that the set D can be described by inequalities of the form

$$x_\ell(y) \le x \le x_u(y), \qquad \text{and } y_\ell \le y \le y_u$$

where y_ℓ, y_u are constants and $x_\ell(y)$, $x_u(y)$ are continuous functions of y on the interval

$$y_\ell \le y \le y_u$$



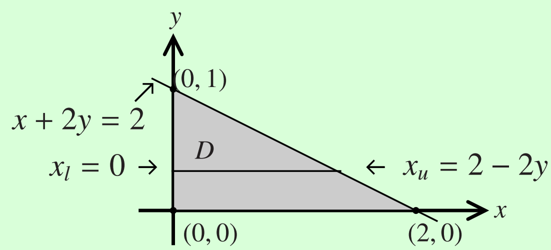
Then, by reversing the roles of x and y in Theorem 1, the double integral $\displaystyle\iint_D f(x,y)\,dA$ can be written as an iterated integral in the order "x first, then y":

$$\iint_D f(x,y)\,dA = \int_{y_\ell}^{y_u} \int_{x_\ell(y)}^{x_u(y)} f(x,y)\,dx\,dy \tag{11.2}$$

EXAMPLE 2 Evaluate the integral in Example 1 by integrating with respect to x first.

Solution: In order to integrate with respect to x first, we describe the set D by $0 \le x \le 2(1-y)$, $0 \le y \le 1$. So, by equation (11.2) we get

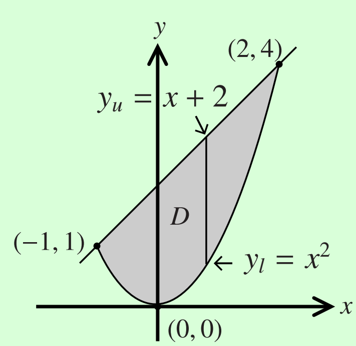
$$\begin{aligned}\iint_D xy\,dA &= \int_{y=0}^{1} \int_{x=0}^{2(1-y)} xy\,dx\,dy \\ &= \int_{y=0}^{1} y\left(\frac{1}{2}x^2\right)\Bigg|_0^{2(1-y)} dy \\ &= 2\int_0^1 y(1-y)^2\,dy = \frac{1}{6}\end{aligned}$$



EXAMPLE 3 Let D be the region bounded by $y = x^2$ and $y = x + 2$. Evaluate $\iint_D (x+2y)\,dA$.

Solution: From the diagram, we observe that the region can be written as $x^2 \le y \le x + 2$ with $-1 \le x \le 2$. Thus,

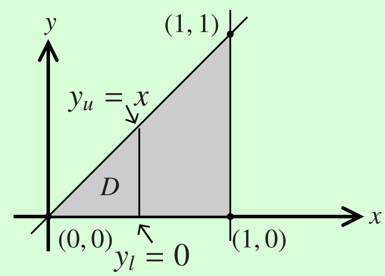
$$\begin{aligned}\iint_D (x+2y)\,dA &= \int_{-1}^{2} \int_{x^2}^{x+2} x + 2y\,dy\,dx \\ &= \int_{-1}^{2} \Big[xy + y^2\Big]_{x^2}^{x+2} dx \\ &= \int_{-1}^{2} (2x^2 + 6x + 4 - x^3 - x^4)\,dx \\ &= \frac{333}{20}\end{aligned}$$



EXAMPLE 4 Let D be the region bounded by the lines $y = 0$, $x = 1$, and $y = x$. Find $\iint_D e^{x^2}\, dA$.

Solution: Although we can easily write the region so that we could integrate with respect either variable first, we see that choosing to integrate with respect to x first would be a bad choice since there is no known anti-derivative of e^{x^2}. Thus, we write the region as $0 \le y \le x$, $0 \le x \le 1$. Then, we get

$$\begin{aligned}\iint_D e^{x^2}\, dA &= \int_0^1 \int_0^x e^{x^2}\, dy\, dx \\ &= \int_0^1 ye^{x^2}\Big|_0^x\, dx \\ &= \int_0^1 xe^{x^2}\, dx \\ &= \frac{1}{2}(e-1)\end{aligned}$$

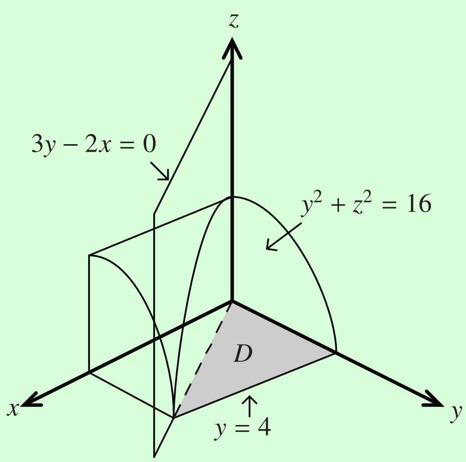


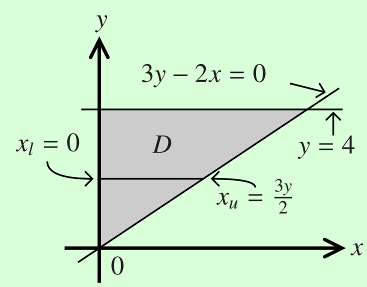
EXAMPLE 5 Find the volume of the solid S in the first octant ($x \ge 0$, $y \ge 0$, $z \ge 0$) bounded by the cylinder $y^2 + z^2 = 16$, and the planes $3y - 2x = 0$, $x = 0$, $z = 0$.

Solution: The cylinder $y^2 + z^2 = 16$ runs parallel to the x-axis (since there is no x-dependence). The plane $3y - 2x = 0$ is vertical (since there is no z-dependence). The solid is described by

$$0 \le z \le \sqrt{16 - y^2} \quad \text{and } (x, y) \in D$$

where D is the region in the xy-plane bounded by $3y - 2x = 0$, $x = 0$, and $y = 4$.





Hence, the volume of the solid is

$$\iint_D \sqrt{16 - y^2}\, dA$$

Observe that we can represent the set D as $0 \le x \le \dfrac{3y}{2}$, and $0 \le y \le 4$. Thus, the volume is

$$\begin{aligned}
\iint\limits_D \sqrt{16-y^2}\, dA &= \int_0^4 \int_0^{3y/2} \sqrt{16-y^2}\, dx\, dy \\
&= \int_0^4 \sqrt{16-y^2}(x)\Big|_0^{3y/2} dy \\
&= \int_0^4 \frac{3}{2} y \sqrt{16-y^2}\, dy \\
&= -\frac{1}{2}(16-y^2)^{3/2}\Big|_0^4 \\
&= 32 \quad \text{cubic units.}
\end{aligned}$$

Observe that the region in Example 5 could have also been represented by $\frac{2x}{3} \le y \le 4$, $0 \le x \le 6$. Hence, we could have applied Theorem 1, instead of using equation (11.2). However, notice that if we had applied Theorem 1 instead, our inner integral would have been

$$\int_{2x/3}^{4} \sqrt{16-y^2}\, dy$$

which would have been more difficult. Thus, when evaluating a double integral

$$\iint\limits_D f(x, y)\, dA$$

one must take into account two factors:

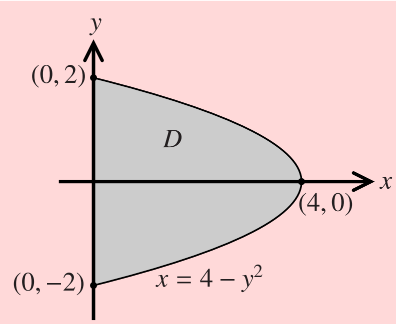
- the shape of the region D.
- the form of the integrand $f(x, y)$.

Either of these factors may make it desirable or even essential to use one order of integration instead of the other.

EXERCISE 1 Describe the set D by inequalities in two ways. Evaluate the double integral

$$\iint\limits_D (x+y)\, dA$$

in two ways.



EXERCISE 2 Let D be the triangular region with vertices $(0,0)$, $(1,1)$, and $(0,2)$. Evaluate

$$\iint_D y\,dA$$

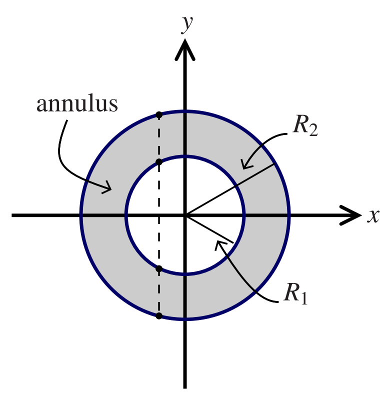
EXERCISE 3 Let D be the triangular region with vertices $(0,0)$, $(0,1)$, and $(1,1)$. Evaluate

$$\iint_D e^{-y^2}\,dA$$

EXERCISE 4 Find the volume of the solid bounded above by the paraboloid $z = 4 - x^2 - y^2$, and below by the rectangle $D = \{(x,y) \mid 0 \le x \le 1, 0 \le y \le 1\}$.

For more complicated regions we may not be able to apply our method above so easily. For example an annulus cannot be described by the usual inequalities since a vertical or a horizontal line may intersect the boundary of D in more than two points. A simple approach to evaluating the double integral

$$\iint_D f(x,y)\,dA$$



where D is the annulus is to let D_1, D_2 denote the discs of radius R_1 and R_2 respectively. Then, by the Decomposition Theorem,

$$\iint_{D_2} f(x,y)\,dA = \iint_{D_1} f(x,y)\,dA + \iint_D f(x,y)\,dA$$

and so the required integral is

$$\iint_D f(x,y)\,dA = \iint_{D_2} f(x,y)\,dA - \iint_{D_1} f(x,y)\,dA$$

Both integrals on the right can be written as iterated integrals in the usual way. However, for this or even more complicated regions, we can often make it simpler by applying a change of variables.

11.3 Change of Variables

In single variable calculus, we learn how a change of variables $x = g(u)$ can be used to transform the integral

$$\int_a^b f(x)\, dx$$

into the integral

$$\int_c^d f(g(u))g'(u)\, du$$

(where c and d are such that $g(c) = a$ and $g(c) = d$) which can sometimes be simpler and easier to evaluate. The change of variables $x = g(u)$ here is technically a function g defined on a suitable domain in $\mathbb{R}$ and subject to certain differentiability and continuity conditions.

We would like to execute a similar process for double integrals. In place of the change of variables $x = g(u)$, what we would like is a *mapping* G given by

$$(x, y) = G(u, v)$$

whose domain is a subset of $\mathbb{R}^2$. Such a mapping can simplify the double integral

$$\iint_{D_{xy}} f(x, y)\, dA$$

either by changing the integrand $f(x, y)$, or by deforming the set D_{xy} in the xy-plane into a simpler shape D_{uv} in the uv-plan (or both).

What remains is to determine the effect of the mapping G on "dA". In the single variable case, the analogous effect is captured by the derivative $g'(u)$, and is usually expressed informally as

$$x = g(u), \quad dx = g'(u)\, du$$

In order to determine the appropriate multivariable analogue, we shall begin by undertaking a study of mappings from $\mathbb{R}^2$ to $\mathbb{R}^2$ in the next chapter.

Chapter 11 Problem Set

1. Show that $\displaystyle\iint_D (ax + by)\, dx\, dy = \frac{1}{3}(a + b)$, where D is the region in the first quadrant bounded by the circle $x^2 + y^2 = 1$ and the lines $x = 0$, $y = 0$; a, b are constants.

2. Find the volume of the solid with height $h(x, y) = xy$ and base D where D is bounded by $y = \frac{1}{2}x$, $y = \sqrt{x}$, $x = 2$ and $x = 4$.

3. Find the volume of the solid with height $h(x, y) = 1+xy$ and base D where D is bounded by $y = x$ and $y = x^2$.

4. Evaluate $\displaystyle\iint_D \sin(x+y)\, dx\, dy$, where D is the triangular region with vertices $(0, 0)$, $(\pi, 0)$ and $\left(\frac{\pi}{2}, \frac{\pi}{2}\right)$.

5. Evaluate the following integrals.

   (a) $\displaystyle\iint_D xy^2\, dA$ where D is the region bounded by $y = x$, $y = 2x$ and $x = 3$.

   (b) $\displaystyle\int_0^1 \int_x^1 y\sqrt{1 - y^3}\, dy\, dx$.

6. Evaluate $\displaystyle\iint_D e^{-y^2}\,dx\,dy$, where D is the triangular region with vertices $(0,0)$, $(0,1)$ and $(1,1)$.

7. For the following iterated integrals sketch the region of integration, and evaluate the integrals by reversing the order of integration:

   (a) $\displaystyle\int_0^1 \left(\int_{x=y}^1 \sin(x^2)\,dx \right) dy$ (b) $\displaystyle\int_0^1 \left(\int_{y=x}^{\sqrt{x}} \frac{\sin y}{y}\,dy \right) dx$

8. Prove that $\displaystyle\iint_D \sin^2(x+y)\,dA \le \iint_D \sin(x+y)\,dA$ where $D = \{(x,y) \mid 0 \le x+y \le \pi \text{ and } 0 \le y \le \pi\}$.

9. Let V denote the volume of the tetrahedron with vertices $(a,0,0)$, $(0,b,0)$, $(0,0,c)$ and $(0,0,0)$, with $a,b,c > 0$. Show that $V = \frac{1}{6}abc$.

10. Let D be the quarter disc in the first quadrant defined by $x^2+y^2 \le 1$. Use the inequality

$$x - \frac{1}{6}x^3 \le \sin x \le x, \text{ for } x \ge 0$$

to show that

$$\frac{14}{45} \le \iint_D \sin x\,dA \le \frac{15}{45}$$

Note: You will not succeed in evaluating this integral exactly.

11. Let D be the unit square $0 \le x \le 1$, $b \le y \le b+1$. Show that

$$\iint_D x^y\,dA = \ln\left(\frac{b+2}{b+1}\right)$$

12. The temperature at points of the disc $x^2+y^2 \le b^2$ is given by

$$f(x,y) = 100 + x^3 - 3xy^2$$

Find the average temperature. At what points of the disc does the temperature equal the average temperature? Give a sketch.

13. Evaluate the iterated integral $\displaystyle\int_1^e \left(\int_{y=0}^{\ln x} \frac{y}{x}\,dy \right) dx$.

14. Evaluate $\displaystyle\iint_D e^{-|x+y|}\,dA$, where $D = \{(x,y) \mid |x| \le 1, |y| \le 1\}$.

15. Evaluate (i) $\displaystyle\iint_D xy\,dA$, (ii) $\displaystyle\iint_D \sin x\,dA$, where D is the unit disc centered at the origin. **Hint:** Don't do much work.

Note: questions 10 and 12 belong in Chapter 13.

Chapter 12

Mappings of $\mathbb{R}^2$ into $\mathbb{R}^2$

So far we have studied scalar-valued functions, that is, functions which map a subset of $\mathbb{R}^2$, or more generally, a subset of $\mathbb{R}^n$ into $\mathbb{R}$. We now extend the ideas of differential calculus to more general functions.

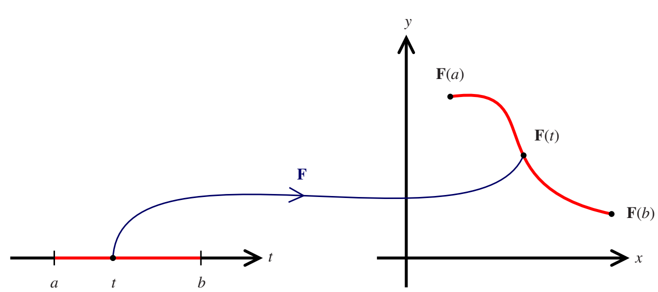
DEFINITION

Vector-Valued Function

A function whose domain is a subset of $\mathbb{R}^n$ and whose codomain is $\mathbb{R}^m$ is called a **vector-valued** function.

You have already worked with the simplest type of vector-valued functions. Consider parametric equations $x = f(t)$, $y = g(t)$ for a curve in $\mathbb{R}^2$: These two scalar equations can be written as a vector equation:

$$(x, y) = F(t) = (f(t), g(t))$$



The function F maps t to $F(t)$, so the domain of F is a subset of $\mathbb{R}$ and its codomain is $\mathbb{R}^2$. Consequently, F is a vector-valued function.

REMARK

While we represent $(f(t), g(t))$ as a point in $\mathbb{R}^2$, remember that it can also be thought of as a position vector.

DEFINITION
Mapping

A vector-valued function whose domain is a subset of $\mathbb{R}^n$ and whose codomain is a subset of $\mathbb{R}^n$ is called a **mapping** (or transformation).

We shall find that linear algebra plays an important role here.

12.1 The Geometry of Mappings

A pair of equations

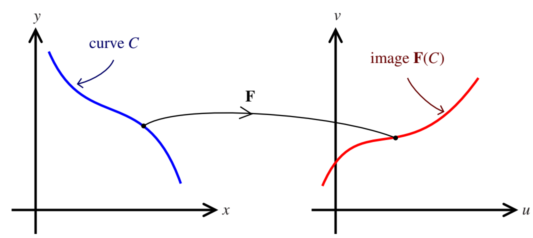
$$\begin{aligned} u &= f(x, y) \\ v &= g(x, y) \end{aligned}$$

associates with each point $(x, y) \in \mathbb{R}^2$ a single point $(u, v) \in \mathbb{R}^2$, and thus defines a vector-valued function

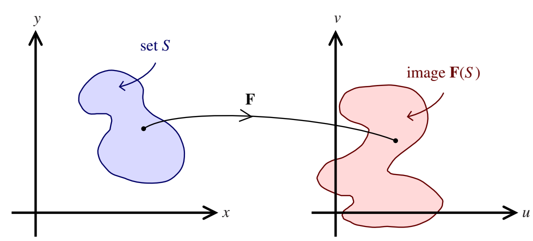
$$(u, v) = F(x, y) = (f(x, y), g(x, y))$$

The scalar functions f and g are called the **component functions** of the mapping. Mappings of $\mathbb{R}^2$ into $\mathbb{R}^2$ (more generally $\mathbb{R}^n$ into $\mathbb{R}^n$) have many applications, such as defining curvilinear coordinate systems (e.g. polar coordinates), and performing a change of variables in multiple integrals (see Sections 13.4 and 14.3). They are used in applied mathematics, in statistics, and in computer graphics for simplifying problems in two or more variables.

In general, if a mapping F from $\mathbb{R}^2$ to $\mathbb{R}^2$ acts on a curve C in its domain, it will determine a curve in its range, denoted by $F(C)$ and called the **image of** C **under** F.



More generally, if a mapping F from $\mathbb{R}^2$ to $\mathbb{R}^2$ acts on any subset S in its domain it will determine a set $F(S)$ in its range, called the **image of** S **under** F.



In order to develop an intuitive geometric understanding of a mapping it is helpful to determine the images of different curves and sets under the mapping. In general, a mapping will deform a given curve or set.

EXAMPLE 1 Consider the mapping defined by $(u, v) = F(x, y) = \left(\frac{1}{2}(x + y), \frac{1}{2}(-x + y)\right)$.

(a) Find the images of the lines $x = k$ and $y = \ell$ under F.

Solution: We are given that $u = \frac{1}{2}(x + y)$ and $v = \frac{1}{2}(-x + y)$. We need to use these equations to convert the equations $x = k$ and $y = \ell$ in terms of u and v.

One way we can do this is to first solve for x and y in terms of u and v.

Observe that we have

$$x = u - v \quad \text{and} \quad y = u + v$$

Thus, a line $x = k$ under the mapping becomes

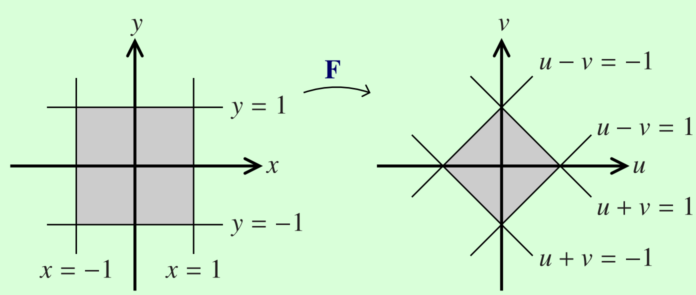
$$u - v = k$$

Similarly, a line $y = \ell$ is transformed into

$$u + v = \ell$$

(b) Find the image of the square $S = \{(x, y) \mid |x| \leq 1, |y| \leq 1\}$ under F.

Solution: To determine the image of S under F, we find the image of each of the boundary lines. In particular, by choosing $k = \pm 1$ and $\ell = \pm 1$, we obtain the images of the sides of the square S.



Observe that the mapping in Example 1 is linear. For any linear mapping, the image of a straight line in the xy-plane is a straight line in the uv-plane. However, we see from the image of S under F that the lines are contracted and rotated by F.

EXERCISE 1 Find the image of the circle $(x - 1)^2 + y^2 = 1$ under the mapping F defined in Example 1.

EXAMPLE 2 Find the image of $D = \{(x, y) \mid -1 \le x \le 3, 0 \le y \le 2\}\}$ under the mapping

$$(u, v) = T(x, y) = (x^2 - y^2, xy)$$

Solution: To determine the image of D under T, we find the image of each of the boundary lines. In this case, it is not so easy to solve for x and y in terms of u and v. We instead substitute the equation of each line directly into the mapping.

For the line $x = -1$, $0 \le y \le 2$, we get

$$u = (-1)^2 - y^2 = 1 - y^2$$
$$v = (-1)y = -y$$

We want equations of curves in the uv-plane, so we eliminate y to obtain

$$u = 1 - (-v)^2 = 1 - v^2$$

Since $v = -y$, the condition $0 \le y \le 2$ gives

$$0 \le -v \le 2 \Rightarrow -2 \le v \le 0$$

For the line $x = 3$, $0 \le y \le 2$, we get $v = 3y$, so

$$u = (3)^2 - y^2 = 9 - y^2 = 9 - \left(\frac{1}{3}v\right)^2 = 9 - \frac{1}{9}v^2$$

with

$$0 \le \frac{1}{3}v \le 2 \Rightarrow 0 \le v \le 6$$

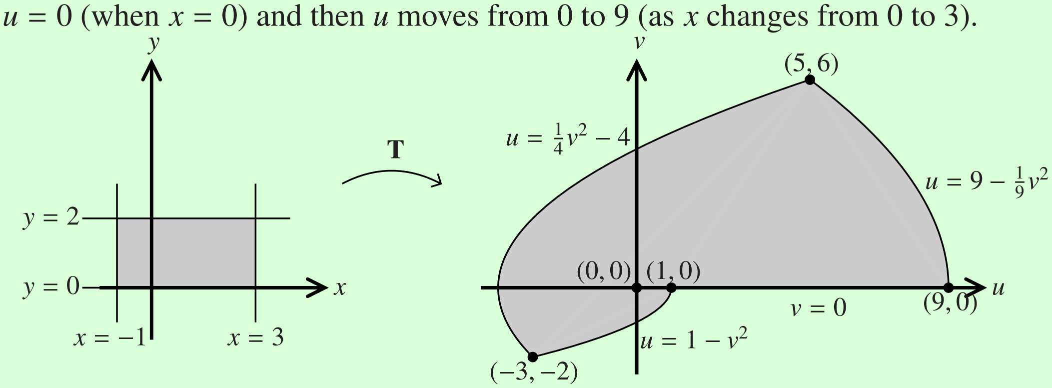
For the line $y = 2$, $-1 \le x \le 3$, we get $v = 2x$, so

$$u = x^2 - 2^2 = x^2 - 4 = \frac{1}{4}v^2 - 4, \quad -2 \le v \le 6$$

For the line $y = 0$, $-1 \le x \le 3$, we get $v = 0$ and

$$u = x^2 - 0^2 = x^2$$

Since x runs from -1 to 3 and $u = x^2$, we get that u starts at 1 (when $x = -1$), moves to $u = 0$ (when $x = 0$) and then u moves from 0 to 9 (as x changes from 0 to 3).



EXAMPLE 3 Find the image of the rectangle

$$R = \left\{(r, \theta) \mid 1 \leq r \leq 2, \quad \frac{\pi}{4} \leq \theta \leq \frac{3\pi}{4}\right\}$$

under the mapping from polar coordinates to Cartesian coordinates defined by

$$(x, y) = F(r, \theta) = (r\cos\theta, r\sin\theta)$$

[Refer to Appendix B for an introduction to polar coordinates.]

Solution: To find the image of the rectangle, we will find the image of each of the boundary lines under F. For the line $r = 1$, $\frac{\pi}{4} \leq \theta \leq \frac{3\pi}{4}$ we get

$$x = \cos\theta, \qquad y = \sin\theta$$

for $\frac{\pi}{4} \leq \theta \leq \frac{3\pi}{4}$. In this case, we don't need to eliminate θ since we recognize these are parametric equations of a circle of radius 1, since they imply

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

Thus, the image is the part of the unit circle with $\frac{\pi}{4} \leq \theta \leq \frac{3\pi}{4}$.

Similarly, we see that the line $r = 2$, $\frac{\pi}{4} \leq \theta \leq \frac{3\pi}{4}$ gives the part of the circle of radius 2 for which $\frac{\pi}{4} \leq \theta \leq \frac{3\pi}{4}$.

The image of a line $\theta = \frac{\pi}{4}$, $1 \leq r \leq 2$ is

$$x = r\cos\frac{\pi}{4} = \frac{1}{\sqrt{2}}r, \qquad y = r\sin\frac{\pi}{4} = \frac{1}{\sqrt{2}}r$$

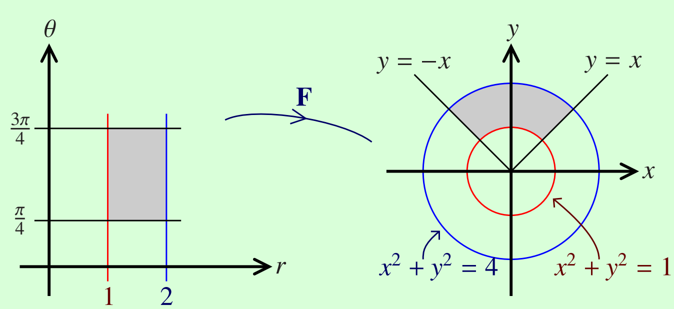
for $1 \leq r \leq 2$. Eliminating r gives $y = x$. Moreover, we have that $r = \sqrt{2}x$ and hence $1 \leq r \leq 2$ gives that x has values from

$$1 \leq \sqrt{2}x \leq 2 \Rightarrow \frac{1}{\sqrt{2}} \leq x \leq \sqrt{2}$$

Similarly, for the line $\theta = \frac{3\pi}{4}$, $1 \leq r \leq 2$ we get

$$x = r\cos\frac{3\pi}{4} = -\frac{1}{\sqrt{2}}r, \qquad y = r\sin\frac{3\pi}{4} = \frac{1}{\sqrt{2}}r$$

for $1 \leq r \leq 2$. Thus, the image is the line $y = -x$ with x values $-\sqrt{2} \leq x \leq -\frac{1}{\sqrt{2}}$.



REMARKS

1. Observe that each of the images are exactly what we would get if we sketched the equations as in Appendix B.
2. The mapping from polar coordinates to Cartesian coordinates is non-linear. The image of a straight line is not necessarily a straight line.

EXERCISE 2 Find the image of the square

$$S = \{(x, y) \mid 1 \leq x \leq 2, 2 \leq y \leq 3\}$$

under the mapping defined by

$$(u, v) = F(x, y) = (xy, y)$$

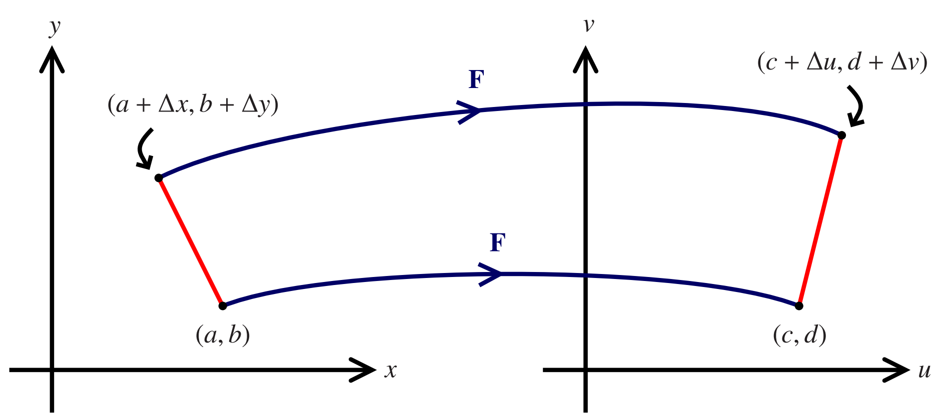
12.2 The Linear Approximation of a Mapping

Consider a mapping F defined by $u = f(x, y)$, $v = g(x, y)$. We assume that F has continuous partial derivatives. By this we mean that the component functions f and g have continuous partial derivatives.

The image of a point (a, b) in the xy-plane is the point (c, d) in the uv-plane, where

$$c = f(a, b), \qquad d = g(a, b)$$

As usual, we want to approximate the image $(c + \Delta u, d + \Delta v)$ of a nearby point $(a + \Delta x, b + \Delta y)$.



We do this by using the linear approximation for $f(x, y)$ and $g(x, y)$ separately. We get

$$\begin{aligned}\Delta u &\approx \frac{\partial f}{\partial x}(a,b)\Delta x + \frac{\partial f}{\partial y}(a,b)\Delta y\\ \Delta v &\approx \frac{\partial g}{\partial x}(a,b)\Delta x + \frac{\partial g}{\partial y}(a,b)\Delta y\end{aligned}$$

for Δx and Δy sufficiently small. This can be written in matrix form as:

$$\begin{bmatrix}\Delta u\\ \Delta v\end{bmatrix} \approx \begin{bmatrix}\frac{\partial f}{\partial x}(a,b) & \frac{\partial f}{\partial y}(a,b)\\ \frac{\partial g}{\partial x}(a,b) & \frac{\partial g}{\partial y}(a,b)\end{bmatrix}\begin{bmatrix}\Delta x\\ \Delta y\end{bmatrix}$$

where the product on the right side of the equation is matrix multiplication.

Observe that this resembles our usual form of the linear approximation where the 2×2 matrix is taking the place of the "derivative". Thus, we make the following definition.

DEFINITION
Derivative Matrix

The **derivative matrix** of a mapping defined by

$$F(x, y) = (f(x, y), g(x, y))$$

is denoted DF and defined by

$$DF = \begin{bmatrix}\frac{\partial f}{\partial x} & \frac{\partial f}{\partial y}\\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y}\end{bmatrix}$$

EXAMPLE 1

Find the derivative matrix of the mapping

$$(u, v) = F(x, y) = (x^2 \sin y, y^2 \cos x)$$

Solution: We have $f(x, y) = x^2 \sin y$ and $g(x, y) = y^2 \cos x$. So,

$$\begin{aligned}DF(x,y) &= \begin{bmatrix}\frac{\partial f}{\partial x} & \frac{\partial f}{\partial y}\\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y}\end{bmatrix}\\ &= \begin{bmatrix}2x\sin y & x^2\cos y\\ -y^2\sin x & 2y\cos x\end{bmatrix}\end{aligned}$$

If we introduce the column vectors

$$\Delta\mathbf{u} = \begin{bmatrix}\Delta u\\ \Delta v\end{bmatrix}, \qquad \Delta\mathbf{x} = \begin{bmatrix}\Delta x\\ \Delta y\end{bmatrix}$$

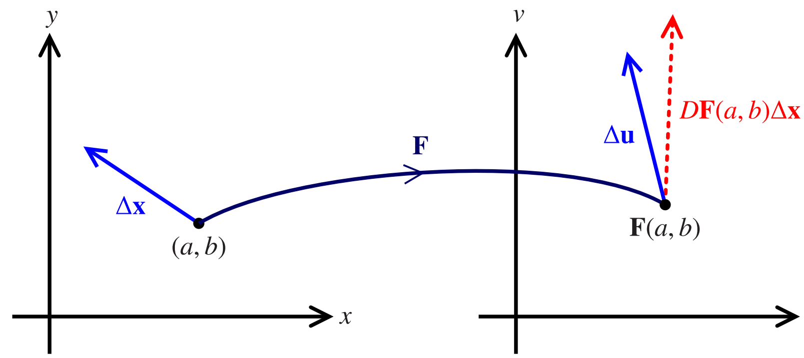
then the **increment form of the linear approximation for mappings** becomes

$$\Delta\mathbf{u} \approx DF(a,b)\Delta\mathbf{x}$$

for $\Delta\mathbf{x}$ sufficiently small. Thus, the **linear approximation for mappings** is

$$F(x,y) \approx F(a,b) + DF(a,b)\Delta\mathbf{x}$$

The geometrical interpretation of the linear approximation for mappings is this: the derivative matrix $DF(a,b)$ acts as a linear mapping on the displacement vector $\Delta\mathbf{x}$ to give an approximation of the image $\Delta\mathbf{u}$ of the displacement under F.



EXAMPLE 2 Consider the mapping defined by

$$(u,v) = F(x,y) = \left(-x + \sqrt{x^2+y^2}, x + \sqrt{x^2+y^2}\right)$$

Use the linear approximation to estimate the image of the point (3.02, 3.99) under F.

Solution: The derivative matrix of F is

$$DF(x,y) = \begin{bmatrix} -1 + \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ 1 + \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \end{bmatrix}$$

As a reference point choose $(3,4)$. Then $F(3,4) = (2,8)$ and

$$DF(3,4) = \begin{bmatrix} -\frac{2}{5} & \frac{4}{5} \\ \frac{8}{5} & \frac{4}{5} \end{bmatrix}$$

The displacement in the uv-plane is approximated by

$$\begin{bmatrix} \Delta u \\ \Delta v \end{bmatrix} \approx DF(3,4)\begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} -\frac{2}{5} & \frac{4}{5} \\ \frac{8}{5} & \frac{4}{5} \end{bmatrix}\begin{bmatrix} 0.02 \\ -0.01 \end{bmatrix} = \begin{bmatrix} -0.016 \\ 0.024 \end{bmatrix}$$

Thus, the linear approximation gives

$$F(3.02, 3.99) \approx (2,8) + (-0.016, 0.024) = (1.984, 8.024)$$

Note: The calculator value is $(1.98405, 8.02405)$.

EXERCISE 1 Consider the mapping defined by

$$(u, v) = F(x, y) = \big(\ln(x + y), \ln(x - y)\big)$$

Approximate the image of the point $(0.95, 0.1)$ under F.

Generalization

A mapping F from $\mathbb{R}^n$ to $\mathbb{R}^m$ is defined by a set of m component functions:

$$\begin{aligned} u_1 &= f_1(x_1, \ldots, x_n) \\ &\vdots \\ u_m &= f_m(x_1, \ldots, x_n) \end{aligned}$$

Or, in vector notation

$$\mathbf{u} = F(\mathbf{x}) = (f_1(\mathbf{x}), \cdots f_m(\mathbf{x})), \quad \mathbf{x} \in \mathbb{R}^n$$

If we assume that F has continuous partial derivatives, then the derivative matrix of F is the $m \times n$ matrix defined by

$$DF(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix}$$

As expected, the linear approximation for F at $\mathbf{a}$ is

$$F(\mathbf{x}) \approx F(\mathbf{a}) + DF(\mathbf{a})\Delta\mathbf{x}$$

where

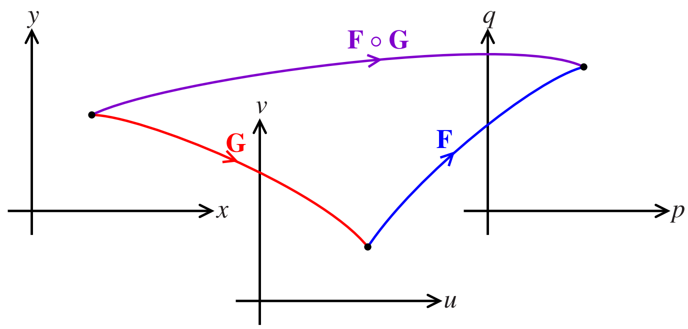
$$\Delta\mathbf{u} = \begin{bmatrix} \Delta u_1 \\ \vdots \\ \Delta u_m \end{bmatrix} \in \mathbb{R}^m, \qquad \Delta\mathbf{x} = \begin{bmatrix} \Delta x_1 \\ \vdots \\ \Delta x_n \end{bmatrix} \in \mathbb{R}^n$$

12.3 Composite Mappings and the Chain Rule

The next step in developing the theory of mappings is to study the composition of two mappings.

Consider successive mappings F and G of $\mathbb{R}^2$ into $\mathbb{R}^2$, defined by

$$F : \begin{cases} p = p(u, v) \\ q = q(u, v) \end{cases} \qquad G : \begin{cases} u = u(x, y) \\ v = v(x, y) \end{cases} \tag{12.1}$$



The composite mapping $F \circ G$, defined by

$$\begin{cases} p = p\big(u(x, y), v(x, y)\big) \\ q = q\big(u(x, y), v(x, y)\big) \end{cases} \tag{12.2}$$

maps the xy-plane directly into the pq-plane.

The question is this: how is the derivative matrix $D(F \circ G)$ of the composite mapping related to the derivative matrices DF and DG of the individual mappings?

The answer is: $D(F \circ G)(x, y)$ is the matrix product of $DF(u, v)$ and $DG(x, y)$, where $(u, v) = G(x, y)$.

We state this formally in the following theorem.

THEOREM 1 **(Chain Rule in Matrix Form)**

Let F and G be mappings from $\mathbb{R}^2$ to $\mathbb{R}^2$. If G has continuous partial derivatives at (x, y) and F has continuous partial derivatives at $(u, v) = G(x, y)$, then the composite mapping $F \circ G$ has continuous partial derivatives at (x, y) and

$$D(F \circ G)(x, y) = DF(u, v)DG(x, y)$$

Proof: Define the component functions for F, G, and $F \circ G$ as in equations (12.1) and (12.2). Then, the chain rule for scalar functions gives

$$\begin{aligned} DF(u, v)DG(x, y) &= \begin{bmatrix} \frac{\partial p}{\partial u} & \frac{\partial p}{\partial v} \\ \frac{\partial q}{\partial u} & \frac{\partial q}{\partial v} \end{bmatrix} \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial p}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial p}{\partial v}\frac{\partial v}{\partial x} & \frac{\partial p}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial p}{\partial v}\frac{\partial v}{\partial y} \\ \frac{\partial q}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial q}{\partial v}\frac{\partial v}{\partial x} & \frac{\partial q}{\partial u}\frac{\partial u}{\partial y} + \frac{\partial q}{\partial v}\frac{\partial v}{\partial y} \end{bmatrix} \\ &= \begin{bmatrix} \frac{\partial p}{\partial x} & \frac{\partial p}{\partial y} \\ \frac{\partial q}{\partial x} & \frac{\partial q}{\partial y} \end{bmatrix} \\ &= D(F \circ G)(x, y) \end{aligned}$$

as required. □

EXAMPLE 1

Consider the mappings G and F defined by

$$(u, v) = G(x, y) = (xy, x + y)$$
$$(p, q) = F(u, v) = (u - v, u^2)$$

Form the composite mapping $F \circ G$ and find the derivative matrices DG, DF, and $D(F \circ G)$. Verify the Chain Rule formula.

Solution: The composite mapping is

$$(p, q) = F(G(x, y)) = F(xy, x + y) = (xy - x - y, x^2y^2)$$

The derivative matrices are:

$$DG(x, y) = \begin{bmatrix} y & x \\ 1 & 1 \end{bmatrix}, \quad DF(u, v) = \begin{bmatrix} 1 & -1 \\ 2u & 0 \end{bmatrix}, \quad D(F \circ G)(x, y) = \begin{bmatrix} y - 1 & x - 1 \\ 2xy^2 & 2x^2y \end{bmatrix}$$

Form the matrix product,

$$\begin{aligned} DF(u, v)DG(x, y) &= \begin{bmatrix} 1 & -1 \\ 2u & 0 \end{bmatrix} \begin{bmatrix} y & x \\ 1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} y - 1 & x - 1 \\ 2uy & 2ux \end{bmatrix} \\ &= \begin{bmatrix} y - 1 & x - 1 \\ 2xy^2 & 2x^2y \end{bmatrix}, \quad \text{on substituting } u = xy \\ &= D(F \circ G)(x, y), \quad \text{as required.} \end{aligned}$$

EXERCISE 1

Consider the mappings defined by

$$F(u, v) = (u^2v, e^{uv-1}), \quad G(x, y) = \left(\sqrt{2x^2 + 2y^2}, 2x + y^2\right)$$

(a) Use the chain rule in matrix form to find the derivative matrix $D(F \circ G)$.
(b) Calculate $D(G \circ F)(1, 1)$.
(c) Use the linear approximation of mappings to approximate the image of $(u, v) = (1.01, 0.98)$ under $G \circ F$.

Chapter 12 Problem Set

1. Consider the following maps. Find the image under T of the square

$$D = \{(x, y) \mid 1 \leq x \leq 2, 2 \leq y \leq 3\}$$

   (a) $T(x, y) = (2x + 3y, x - y)$
   (b) $T(x, y) = (xy, x^2 - y^2)$
   (c) $T(x, y) = (x \cos \frac{1}{3}\pi y, x \sin \frac{1}{3}\pi y)$
   (d) $T(x, y) = (e^{x+y}, e^{x-y})$

2. Consider $(u, v) = F(x, y) = (y + e^x, e^x - y)$.
   (a) Find and sketch the image of the square with vertices $(0, 0)$, $(0, 1)$, $(1, 1)$, and $(1, 0)$ under F.
   (b) Use the linear approximation for mappings to approximate the image of $(x, y) = (0.01, 0.02)$ under F.

3. Find the image of $D = \{(x, y) \in \mathbb{R}^2 | -1 \leq x \leq 2, 0 \leq y \leq 2\}$ under $T(x, y) = (x + 2y, 3x - y)$.

4. Find the image of $D = \{(x, y) \in \mathbb{R}^2 | 1 \leq x \leq 2, 0 \leq y \leq 2\}$ under $T(x, y) = (x^2 + y^2, x^2 - y^2)$.

5. Find the image of the annulus $4 \leq x^2 + y^2 \leq 16$ under the map defined by

$$(u, v) = F(x, y) = \left(\frac{x}{x^2 + y^2}, \frac{y}{x^2 + y^2}\right)$$

6. Use the linear approximation in matrix form to find the approximate image of the point (3.1, 3.9) under the map defined by

$$(u, v) = F(x, y) = \left(\sqrt{x^2 + y^2}, \frac{x}{\sqrt{x^2 + y^2}}\right)$$

7. Let $(p, q) = F(u, v) = \left(v \cos(uv + 2), v\sqrt{u^2 + 5}\right)$ and $(u, v) = G(x, y) = (xy - xy^2, xe^{xy-2})$.
   (a) Use the chain rule in matrix form to find $D(F \circ G)(1, 2)$.
   (b) Use the linear approximation for mappings to approximate the image of $(x, y) = (1.1, 1.9)$ under $F \circ G$.

8. Consider the maps F and G defined by

$$F(u, v) = (e^{u+v}, e^{u-v}), \qquad G(x, y) = (xy, x^2 - y^2)$$

   (a) Calculate the composite map $F \circ G$ and the derivative matrix $D(F \circ G)(1, 1)$.
   (b) Verify your answer for $D(F \circ G)(1, 1)$ by using the Chain Rule in matrix form.
   (c) Calculate $D(G \circ F)(1, 1)$.

9. Consider the maps F and G defined by

$$F(u, v) = (v + u^2, u), \qquad G(x, y) = (e^x y, 2e^{-x} y)$$

   State the Chain Rule in matrix form, and use it to calculate the derivative $D(F \circ G)(0, 1)$ of the composite map.

10. Let $(u, v) = F(x, y) = \left(x \ln(y - x^4), (2 + \frac{y}{x})^{3/2}\right)$. Suppose that $G(u, v)$ has continuous partial derivatives with $G(0, 8) = (1, -1)$ and $DG(0, 8) = \begin{bmatrix} -2 & 1 \\ -4 & 3 \end{bmatrix}$. Use the linear approximation to approximate $(G \circ F)(0.9, 2.1)$.

11. (a) Let $(u, v) = F(x, y)$ be a map of the xy-plane into the uv-plane. Consider a smooth curve $(x(t), y(t))$ in the xy-plane. Suppose that F maps this curve into the curve $(u(t), v(t))$ in the uv-plane. Show that the tangent vectors are related by the derivative matrix according to

$$\begin{bmatrix} u'(t) \\ v'(t) \end{bmatrix} = DF(x(t), y(t)) \begin{bmatrix} x'(t) \\ y'(t) \end{bmatrix}$$

   (b) Consider the map $(u, v) = (xy, x^2 - y^2)$. Find the image of the curve $(x, y) = (t, t^2)$, $t \geq 0$ under this map, and sketch both curves. Calculate the tangent vectors to the curves, and verify the formula that you derived in part (a).

12. * Sketch the image of the square

$$D = \{(x, y) \mid 1 \leq x \leq 2, 2 \leq y \leq 3\}$$

   under the map

$$T(x, y) = \left(x \cos\left(\frac{\pi}{3} xy\right), x \sin\left(\frac{\pi}{3} xy\right)\right)$$

Chapter 13

Jacobians and the Change of Variables Theorem

13.1 The Inverse Mapping Theorem

Our goal now is to find a condition which will guarantee that a mapping $(u, v) = F(x, y)$ has an inverse. We start by defining inverse mappings in the expected way.

DEFINITION

Invertible Mapping

Inverse Mapping

Let F be a mapping from a set D_{xy} onto a set D_{uv}. If there exists a mapping F^{-1}, called the **inverse of** F which maps D_{uv} onto D_{xy} such that

$$(x, y) = F^{-1}(u, v) \quad \text{if and only if} \quad (u, v) = F(x, y)$$

then F is said to be **invertible** on D_{xy}.

As usual, we have

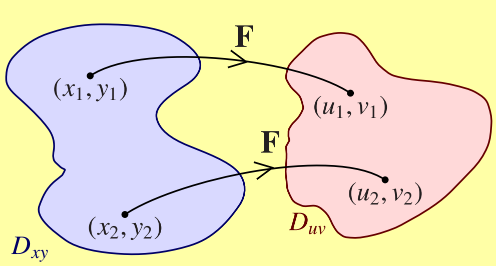
$$\begin{aligned} (F^{-1} \circ F)(x, y) &= (x, y) \qquad \text{for all } (x, y) \in D_{xy} \\ (F \circ F^{-1})(u, v) &= (u, v) \qquad \text{for all } (u, v) \in D_{uv} \end{aligned} \tag{13.1}$$

Recall that a function being invertible is related to it being one-to-one.

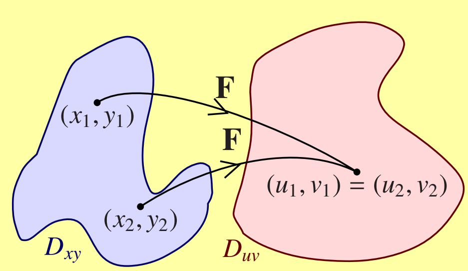
DEFINITION

One-to-One

A mapping F from $\mathbb{R}^2$ to $\mathbb{R}^2$ is said to be **one-to-one** on a set D_{xy} if and only if $F(a, b) = F(c, d)$ implies $(a, b) = (c, d)$, for all $(a, b), (c, d) \in D_{xy}$.



F is one-to-one



F is not one-to-one

THEOREM 1 Let F be a mapping from a set D_{xy} onto a set D_{uv}. If F is one-to-one on D_{xy}, then F is invertible on D_{xy}.

Now, recall from Calculus 1 that if $f'(x) > 0$ for all $x \in [a, b]$, then f is one-to-one on $[a, b]$ and hence has an inverse on $[a, b]$. Thus, for a mapping F, it makes sense to investigate the relation between the derivative matrix DF of F and F being invertible. We start with the following theorem.

THEOREM 2 Consider a mapping F which maps D_{xy} onto D_{uv}. If F has continuous partial derivatives at $\mathbf{x} \in D_{xy}$ and there exists an inverse mapping F^{-1} of F which has continuous partial derivatives at $\mathbf{u} = F(\mathbf{x}) \in D_{uv}$, then

$$DF^{-1}(\mathbf{u})DF(\mathbf{x}) = I$$

Proof: By the Chain Rule in Matrix Form we get

$$DF^{-1}(\mathbf{u})DF(\mathbf{x}) = D(F^{-1} \circ F)(\mathbf{x})$$

Then, by equation (13.1) we have

$$D(F^{-1} \circ F)(\mathbf{x}) = D\mathbf{x} = \begin{bmatrix} \frac{\partial x}{\partial x} & \frac{\partial x}{\partial y} \\ \frac{\partial y}{\partial x} & \frac{\partial y}{\partial y} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

as required. □

EXAMPLE 1 Consider the mapping defined by

$$(u, v) = F(x, y) = (y + x^2, x)$$

Solve for the inverse mapping F^{-1}. Find the derivative matrices DF and DF^{-1} and verify that $DF^{-1}(u, v)$ is the matrix inverse of $DF(x, y)$.

Solution: The inverse mapping is obtained by solving

$$u = y + x^2, \quad v = x$$

for x and y. We obtain

$$x = v, \qquad y = u - v^2$$

Hence, the inverse mapping is

$$(x, y) = F^{-1}(u, v) = (v, u - v^2)$$

The derivative matrices are:

$$DF(x, y) = \begin{bmatrix} 2x & 1 \\ 1 & 0 \end{bmatrix}, \quad DF^{-1}(u, v) = \begin{bmatrix} 0 & 1 \\ 1 & -2v \end{bmatrix}$$

Form the matrix product,

$$\begin{aligned} DF^{-1}(u, v)DF(x, y) &= \begin{bmatrix} 0 & 1 \\ 1 & -2v \end{bmatrix}\begin{bmatrix} 2x & 1 \\ 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 2x - 2v & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{on substituting } v = x. \end{aligned}$$

REMARK

The fact that we could solve and obtain a unique solution for x and y in the preceding example proves that F has an inverse mapping on $\mathbb{R}^2$. It is only in simple examples that one can carry out this step. Hence it is useful to develop a test to determine if a mapping F has an inverse mapping.

The determinant of the derivative matrix plays an important role in the study of mappings and in their application to multiple integrals.

DEFINITION

Jacobian

The **Jacobian** of a mapping

$$(u, v) = F(x, y) = (u(x, y), v(x, y))$$

is denoted $\dfrac{\partial(u, v)}{\partial(x, y)}$, and is defined by

$$\frac{\partial(u, v)}{\partial(x, y)} = \det[DF(x, y)] = \det \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix}$$

REMARK

The Jacobian is the key missing ingredient that we needed to complete our change of variables procedure outlined in Section 11.3. See Section 13.4.

EXERCISE 1 Calculate the Jacobian $\dfrac{\partial(x,y)}{\partial(r,\theta)}$ of the mapping F given by

$$(x,y) = F(r,\theta) = (r\cos\theta, r\sin\theta)$$

One can interpret Theorem 2 as asserting that if a mapping F is invertible, then its derivative matrix $DF(x,y)$ is invertible, and its inverse matrix is the derivative matrix $DF^{-1}(u,v)$ of the inverse map. Recall from linear algebra that a square matrix has an inverse matrix if and only if its determinant is non-zero. Thus, it follows from Theorem 2 that if a mapping F has an inverse mapping F^{-1} (and both mappings have continuous partial derivatives), then the Jacobian of F is non-zero. This is stated as a corollary to Theorem 2.

COROLLARY 3 Consider a mapping defined by

$$(u,v) = F(x,y) = (f(x,y), g(x,y))$$

which maps a subset D_{xy} onto a subset D_{uv}. Suppose that f and g have continuous partials on D_{xy}. If F has an inverse mapping F^{-1}, with continuous partials on D_{uv}, then the Jacobian of F is non-zero:

$$\frac{\partial(u,v)}{\partial(x,y)} \neq 0, \qquad \text{on } D_{xy}$$

REMARK

The notation $\dfrac{\partial(u,v)}{\partial(x,y)}$ for the Jacobian reminds one which partial derivatives have to be calculated. Thus, if F maps $(x,y) \to (u,v)$ and is one-to-one, then the inverse mapping F^{-1} maps $(u,v) \to (x,y)$, and the Jacobian of the inverse mapping is denoted by

$$\frac{\partial(x,y)}{\partial(u,v)} = \det[F^{-1}(u,v)] = \det\begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix}$$

Recall from linear algebra that $\det(AB) = \det A \det B$ for all $n \times n$ matrices A, B. Thus, we can deduce from Theorem 2 a simple relationship between the Jacobian of a mapping and the Jacobian of the inverse mapping. We state this as another corollary to Theorem 2.

COROLLARY 4 **(Inverse Property of the Jacobian)**

If the hypotheses of Theorem 2 hold, then

$$\frac{\partial(x,y)}{\partial(u,v)} = \frac{1}{\dfrac{\partial(u,v)}{\partial(x,y)}}$$

Proof: By Theorem 2,

$$I = DF^{-1}(u,v)DF(x,y)I$$

Taking the determinant of this equation gives

$$\begin{aligned} \det I &= \det(DF^{-1}(u,v)DF(x,y)) \\ 1 &= \det(DF^{-1}(u,v))\det(DF(x,y)) \end{aligned}$$

Thus, by definition of the Jacobian,

$$1 = \frac{\partial(x,y)}{\partial(u,v)}\frac{\partial(u,v)}{\partial(x,y)}$$

Since $DF(x,y)$ is invertible, we have $\frac{\partial(u,v)}{\partial(x,y)} \neq 0$. Therefore, we get

$$\frac{\partial(x,y)}{\partial(u,v)} = \frac{1}{\dfrac{\partial(u,v)}{\partial(x,y)}}$$

$\square$

Since we are interested in being able to test whether or not F^{-1} exists, we ask: does Corollary 3 admit a converse? i.e. does $\dfrac{\partial(u,v)}{\partial(x,y)} \neq 0$ on D_{xy} imply that F^{-1} exists? Unfortunately NO, unless we formulate the question more carefully. The following example shows what can go wrong.

EXAMPLE 2 Consider the mapping defined by

$$(u,v) = F(x,y) = (e^x \cos y, e^x \sin y)$$

Show that $\dfrac{\partial(u,v)}{\partial(x,y)} \neq 0$ on $\mathbb{R}^2$, but that F^{-1} does not exist on $\mathbb{R}^2$.

Solution: Observe that

$$\frac{\partial(u,v)}{\partial(x,y)} = e^{2x} > 0 \quad \text{for all } (x,y) \in \mathbb{R}^2$$

However, F is not one-to-one on $\mathbb{R}^2$, since, for example

$$F(0,0) = F(0,2\pi) = (1,0)$$

Thus, F^{-1} does not exist on $\mathbb{R}^2$.

The reason the mapping in Example 2 is not invertible is because of the periodic behavior of $\sin y$ and $\cos y$. However, we know we can create inverse functions for these by restricting their domain to a neighborhood where they are one-to-one. Similarly, in Example 2, if we restrict the domain to a neighborhood $N(0,0)$ of radius less than 2π, it will be possible to solve uniquely for x and y in terms of u and v; in particular, an inverse mapping does exist. We can generalize this into the following theorem.

THEOREM 5 **(Inverse Mapping Theorem)**

If a mapping $(u,v) = F(x,y)$ has continuous partial derivatives in some neighborhood of (a,b) and $\dfrac{\partial(u,v)}{\partial(x,y)} \neq 0$ at (a,b), then there is a neighborhood of (a,b) in which F has an inverse mapping $(x,y) = F^{-1}(u,v)$ which has continuous partial derivatives.

The proof is beyond the scope of this course.

EXAMPLE 3 Consider the mapping defined by

$$(u,v) = F(x,y) = (xy - x^2, x + y)$$

Show that F has an inverse mapping in a neighborhood of $(1,-2)$.

Solution: The Jacobian of F is

$$\frac{\partial(u,v)}{\partial(x,y)} = \det\begin{bmatrix} y-2x & x \\ 1 & 1 \end{bmatrix} = y - 3x$$

Hence at $(x,y) = (1,-2)$, the Jacobian is non-zero. Clearly the partial derivatives of F are continuous by the Continuity Theorems. Thus, by the Inverse Mapping Theorem, there is a neighborhood of $(1,-2)$ in which F has an inverse mapping.

EXERCISE 2 Referring to Example 3, show that the inverse mapping is given by

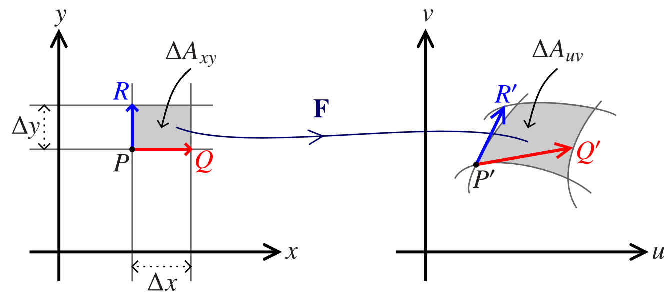
$$(x,y) = F^{-1}(u,v) = \left(\frac{1}{4}(v + \sqrt{v^2 - 8u}), \frac{1}{4}(3v - \sqrt{v^2 - 8u})\right)$$

13.2 Geometrical Interpretation of the Jacobian

In this section, we explain the geometrical interpretation of the Jacobian of a mapping. This interpretation is based on the following result from linear algebra. The area of a parallelogram defined by the vectors $\begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$ and $\begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$ is given by

$$\text{Area} = \left|\det\begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \end{bmatrix}\right| \tag{13.2}$$

We calculate the area of the image in the uv-plane, of a small rectangle in the xy-plane under a mapping F.



We approximate the image of the rectangle defined by the vectors $\overrightarrow{PQ}$ and $\overrightarrow{PR}$ as a parallelogram defined by the vectors $\overrightarrow{P'Q'}$ and $\overrightarrow{P'R'}$, and use the linear approximation to approximate $\overrightarrow{P'Q'}$ and $\overrightarrow{P'R'}$.

Since $\overrightarrow{PQ} = \begin{bmatrix} \Delta x \\ 0 \end{bmatrix}$ and $\overrightarrow{PR} = \begin{bmatrix} 0 \\ \Delta y \end{bmatrix}$, we obtain

$$\overrightarrow{P'Q'} \approx \begin{bmatrix} u_x & u_y \\ v_x & v_y \end{bmatrix} \begin{bmatrix} \Delta x \\ 0 \end{bmatrix} = \begin{bmatrix} u_x \Delta x \\ v_x \Delta x \end{bmatrix}$$

$$\overrightarrow{P'R'} \approx \begin{bmatrix} u_x & u_y \\ v_x & v_y \end{bmatrix} \begin{bmatrix} 0 \\ \Delta y \end{bmatrix} = \begin{bmatrix} u_y \Delta y \\ v_y \Delta y \end{bmatrix}$$

for Δx and Δy sufficiently small. Note that the partial derivatives are evaluated at P. We have

$$\Delta A_{xy} = \Delta x \Delta y$$

and so, by (13.2),

$$\Delta A_{uv} \approx \left| \det \begin{bmatrix} u_x \Delta x & u_y \Delta y \\ v_x \Delta x & v_y \Delta y \end{bmatrix} \right| = \left| \det \begin{bmatrix} u_x & u_y \\ v_x & v_y \end{bmatrix} \right| \Delta x \Delta y$$

since Δx and Δy are positive. Thus, by definition of the Jacobian,

$$\Delta A_{uv} \approx \left| \frac{\partial(u, v)}{\partial(x, y)} \right| \Delta A_{xy} \tag{13.3}$$

where the Jacobian is evaluated at P.

In words, the Jacobian of a mapping F describes the extent to which F increases or decreases areas. We can think of the Jacobian of F as a magnification factor for (very small) areas that are mapped by F. Keep in mind that the basic relation (13.3) is an approximation, which is valid only for small areas, and which becomes increasingly accurate as Δx and Δy tend to zero.

EXAMPLE 1 Calculate the approximate area of the image of a small rectangle of area $\Delta x \Delta y$, located at the point $(3, 4)$, under the mapping F defined by

$$(u, v) = F(x, y) = \left(-x + \sqrt{x^2 + y^2}, \quad x + \sqrt{x^2 + y^2}\right)$$

Solution: Differentiation and evaluation at $(3, 4)$ gives the derivative matrix at $(3, 4)$:

$$DF(3, 4) = \begin{bmatrix} -\frac{2}{5} & \frac{4}{5} \\ \frac{8}{5} & \frac{4}{5} \end{bmatrix}$$

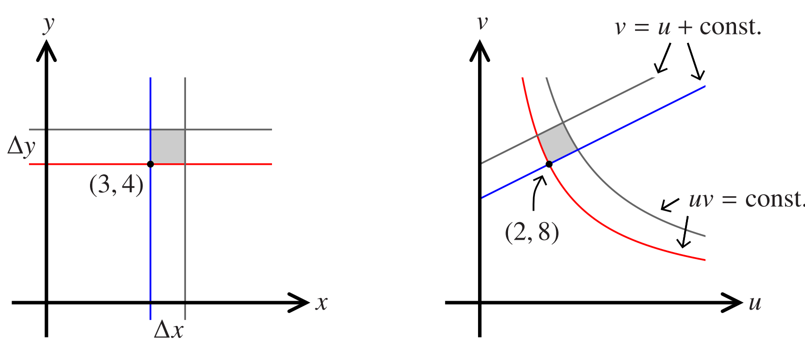
At $(3, 4)$ the Jacobian is

$$\frac{\partial(u, v)}{\partial(x, y)} = \det \begin{bmatrix} -\frac{2}{5} & \frac{4}{5} \\ \frac{8}{5} & \frac{4}{5} \end{bmatrix} = -\frac{8}{5}$$

Therefore, the area of the image is approximately

$$\Delta A_{uv} \approx \frac{8}{5} \Delta A_{xy}$$

We can use a diagram to demonstrate what is happening geometrically in the example.

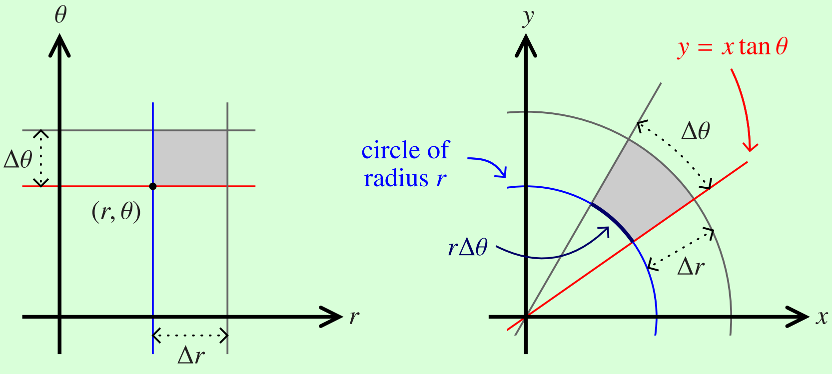


EXAMPLE 2 Consider the mapping F defined by

$$(x, y) = F(r, \theta) = (r \cos \theta, r \sin \theta)$$

Find the image in the xy-plane, of a rectangle in the $r\theta$-plane, and verify directly that the Jacobian gives the magnification factor for area.

Solution: Using what we did in Example 12.1.3, we find that the images of the lines $r = k$ and $\theta = \ell$ are the circles $x^2 + y^2 = k^2$ and the lines $y = x \tan \theta$.



The area of the rectangle in the $r\theta$-plane is

$$\Delta A_{r\theta} = \Delta r \Delta\theta$$

The image of this rectangle in the xy-plane can be approximated by a rectangle with sides of length $r\Delta\theta$ and Δr, for Δr and $\Delta\theta$ sufficiently small. So,

$$\Delta A_{xy} \approx r\Delta r\Delta\theta = r\Delta A_{r\theta}$$

However, the Jacobian of the mapping is

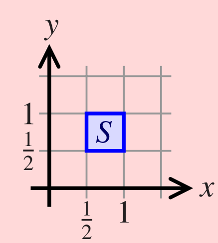
$$\frac{\partial(x, y)}{\partial(r, \theta)} = r > 0$$

(Exercise 13.1.1). Consequently,

$$\Delta A_{xy} \approx \left|\frac{\partial(x, y)}{\partial(r, \theta)}\right| \Delta A_{r\theta}$$

which verifies the area transformation formula (13.3).

EXERCISE 1 Let $F(x, y) = (x^2y, -xy)$ and let S be the square pictured in the diagram. Will the image of S under F have more or less area? Explain your answer.



REMARK

For a linear mapping $(u, v) = F(x, y) = (ax+by, cx+dy)$ where a, b, c, d are constants, the derivative matrix is

$$DF(x, y) = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

and thus the linear approximation is exact by Taylor's Theorem since all second partials are zero. Therefore, for a linear mapping the approximation (13.3) becomes an exact relation.

EXERCISE 2 Show that the linear mapping $(u, v) = F(x, y) = (x + 2y, x + y)$ preserves areas. Illustrate the action of the mapping by finding the image of the square with vertices $(0, 0)$, $(0, 1)$, $(1, 0)$ and $(1, 1)$.

EXERCISE 3 Use the Jacobian to verify the well-known result that any linear mapping F which is a rotation,

$$(u, v) = F(x, y) = (x \cos\theta + y \sin\theta, -x \sin\theta + y \cos\theta)$$

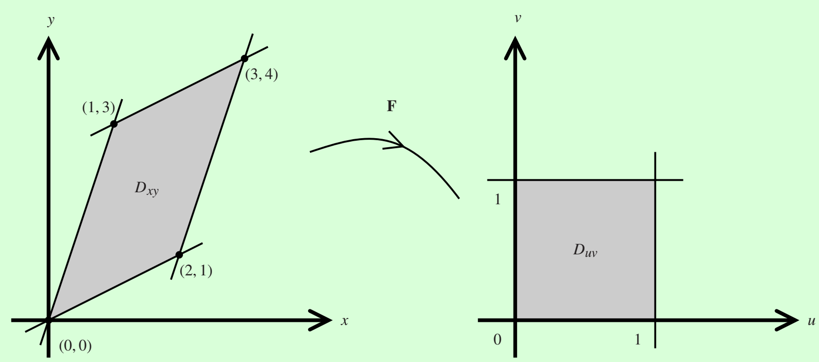
where θ is a constant, preserves areas.

13.3 Constructing Mappings

When performing change of variables in double and triple integrals, it will be very important to be able to invent an invertible mapping which transforms one region to another, simpler region. We demonstrate this with some examples.

EXAMPLE 1 Find a linear mapping F which will transform the parallelogram with vertices $(0, 0)$, $(2, 1)$, $(3, 4)$ and $(1, 3)$ in the xy-plane into the unit square $0 \le u \le 1$, $0 \le v \le 1$ in the uv-plane. Calculate the Jacobian of F and hence find the area of the parallelogram.

Solution:



Solution: The lines bounding D_{xy} are $2y - x = 0$, $2y - x = 5$, $3x - y = 0$, and $3x - y = 5$. We recall from chapter 12, that when performing a mapping, we substituted the equations of each line into the component functions. Thus, we want to pick component functions $u = f(x, y)$, $v = g(x, y)$, so that the image of the lines are $u = 0$, $u = 1$, $v = 0$, and $v = 1$ respectively. Observe, that the bounding lines come in pairs. To get the first pair to have images $u = 0$ and $u = 1$, we see that we can take $u = \frac{2y-x}{5}$. For the second pair to have images $v = 0$ and $v = 1$ we take $v = \frac{3x-y}{5}$. Thus, the desired mapping is

$$(u, v) = F(x, y) = \left(\frac{2y - x}{5}, \frac{3x - y}{5} \right)$$

The Jacobian is

$$\frac{\partial(u,v)}{\partial(x,y)} = \det\begin{bmatrix} -\frac{1}{5} & \frac{2}{5} \\ \frac{3}{5} & -\frac{1}{5} \end{bmatrix} = -\frac{1}{5}$$

Since the mapping is linear, we have the exact relation

$$A_{uv} = \left|\frac{\partial(u,v)}{\partial(x,y)}\right| A_{xy} = \frac{1}{5}A_{xy}$$

Hence, the area of the parallelogram is 5 square units.

EXAMPLE 2 Find a linear mapping which transforms the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ into the unit circle $u^2 + v^2 = 1$.

Solution: We want to pick $u = f(x,y)$ and $v = g(x,y)$, such that we turn $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ into $u^2 + v^2 = 1$. If we write the ellipse as $\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1$, then it is clear that we want to take $u = \frac{x}{a}$ and $v = \frac{y}{b}$. Hence, the desired mapping is

$$(u,v) = F(x,y) = \left(\frac{x}{a}, \frac{y}{b}\right)$$

EXERCISE 1 Find a linear mapping F which transforms the ellipse $3x^2 + 2xy + y^2 = 4$ into the circle $u^2 + v^2 = 4$.

EXAMPLE 3 Find an invertible mapping which will transform the region D_{xy} in the first quadrant bounded by the hyperbola $xy = 1$, $xy = 3$, $x^2 - y^2 = 2$, $x^2 - y^2 = 4$ into a square in the uv-plane.

Solution: We again see that we have pairs of equations. Thus, if we take $u = xy$ and $v = x^2 - y^2$ we see that the images of the hyperbola $xy = 1$, $xy = 3$, $x^2 - y^2 = 2$, $x^2 - y^2 = 4$ are $u = 1$, $u = 3$, $v = 2$, $v = 4$. Hence, the mapping

$$(u,v) = F(x,y) = (xy, x^2 - y^2)$$

gives the desired transformation. Observe that it would be difficult to solve for the inverse explicitly, however, we can at least show that the mapping is locally invertible by applying the Inverse Mapping Theorem. The Jacobian of F is

$$\det DF(x,y) = \det\begin{bmatrix} y & x \\ 2x & -2y \end{bmatrix} = -2x^2 - 2y^2$$

which is non-zero on the region D_{xy} and F has continuous partial derivatives, so F is invertible in a neighborhood of every point in D_{xy} by the Inverse Mapping Theorem.

REMARK

In Example 3, the solution does not actually prove that the mapping is invertible on the entire region. In practice, we often assume that "invertible in a neighborhood of each point" implies "invertible over the entire region", but this is not always true. (For example, $F(r,\theta) = (r\cos\theta, r\sin\theta)$ on $1 \le r \le 2$, $0 \le \theta \le 4\pi$ has nonzero Jacobian at each point but is not one-to-one due to periodicity.)

EXERCISE 2 Find an invertible mapping which will transform the region D_{xyz} in the first octant bound by $xy = 1$, $xy = 3$, $xz = 1$, $xz = 3$, $yz = 2$, and $yz = 4$ into a cube in the uvw-space.

13.4 The Change of Variables Theorem for Double Integrals

We are now in a position to describe the **change of variables** process for double integrals that was hinted at in Section 11.3.

Recall that our objective was to examine the effect of a mapping

$$(x, y) = G(u, v) \tag{13.4}$$

on a double integral

$$\iint_{D_{xy}} f(x, y)\, dA$$

The mapping changes the integrand from $f(x, y)$ to $f(G(u, v))$ and it transforms the region of integration from D_{xy} in the xy-plane to the region D_{uv} in the uv-plane.

In this type of calculation it is convenient to replace the symbol "dA" in the double integral by "$dx\, dy$" if one is working in the xy-plane, and by "$du\, dv$" if one is working in the uv-plane.

In order to derive the change of variables formula for double integrals, we need the formula which describes how areas are related under a mapping G given by (13.4). The geometric interpretation of the Jacobian gives us

$$\Delta A_{xy} \approx \left|\frac{\partial(x, y)}{\partial(u, v)}\right| \Delta A_{uv} \tag{13.5}$$

for Δu, Δv sufficiently small where the Jacobian $\dfrac{\partial(x, y)}{\partial(u, v)}$ is evaluated at a point in the region. Notice that we have interchanged the roles of (x, y) and (u, v) in equations (13.4) and (13.5), as compared to Section 13.2.

THEOREM 1 **(Change of Variables Theorem)**

Let each of D_{uv} and D_{xy} be a closed bounded set whose boundary is a piecewise-smooth closed curve. Let

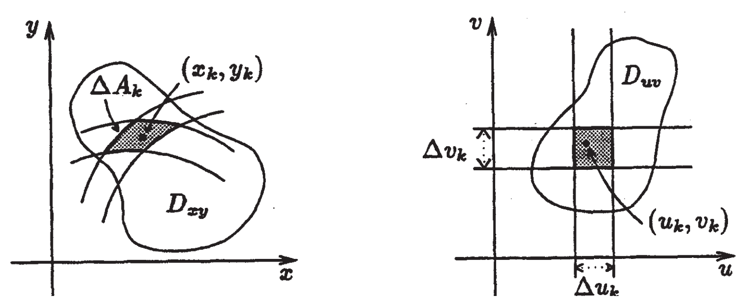
$$(x, y) = G(u, v) = (g(u, v), h(u, v))$$

be a one-to-one mapping of D_{uv} onto D_{xy}, with $g, h \in C^1$, and $\dfrac{\partial(x, y)}{\partial(u, v)} \neq 0$ except for possibly on a finite collection of piecewise-smooth curves in D_{uv}. If $f(x, y)$ is continuous on D_{xy}, then

$$\iint_{D_{xy}} f(x, y)\, dx\, dy = \iint_{D_{uv}} f\Big(g(u, v), h(u, v)\Big) \left|\frac{\partial(x, y)}{\partial(u, v)}\right| du\, dv$$

While a proof of this theorem is beyond the scope of this course, we can make the result plausible by the following argument:

Consider a partition P of D_{uv} into rectangles, by means of straight lines parallel to the coordinate axes. The images of these lines under the given transformation will in general be two families of curves which will define a partition P^* of D_{xy} into elements of area which are approximately parallelograms. We can use this partition, instead of a rectangular partition, to define $\displaystyle\iint_{D_{xy}} f(x, y)\, dx\, dy$.



Thus,

$$\begin{aligned}
\iint_{D_{xy}} f(x, y)\, dx\, dy &= \lim_{\Delta P^* \to 0} \sum_{i=1}^{n} f(x_i, y_i)\Delta A_i \\
&= \lim_{\Delta P^* \to 0} \sum_{i=1}^{n} f\Big(g(u_i, v_i), h(u_i, v_i)\Big) \left|\frac{\partial(x, y)}{\partial(u, v)}\right|_{(u_i, v_i)} \Delta A_i \\
&= \iint_{D_{uv}} f\Big(g(u, v), h(u, v)\Big) \left|\frac{\partial(x, y)}{\partial(u, v)}\right| dA
\end{aligned}$$

by using the definition of double integral relative to the rectangular partition of D_{uv}.

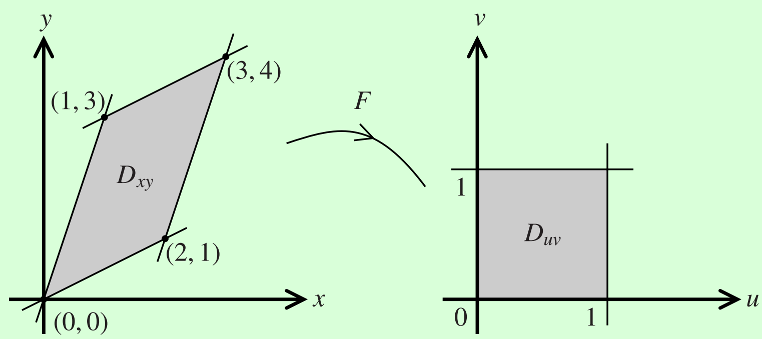
The lack of rigor occurs when we use the approximation (13.5).

EXAMPLE 1 Evaluate $\displaystyle\iint_{D_{xy}} (x+y)\, dA$, where D_{xy} is the set bounded by the parallelogram with vertices $(0,0)$, $(2,1)$, $(1,3)$, and $(3,4)$.

Solution: In Example 13.3.1, we found that the mapping

$$(u,v) = F(x,y) = \left(\frac{1}{5}(2y-x), \frac{1}{5}(3x-y)\right)$$

maps D_{xy} onto D_{uv}, the unit square in the uv-plane.



The Jacobian of F is

$$\frac{\partial(u,v)}{\partial(x,y)} = -\frac{1}{5}$$

Observe that our mapping F maps D_{xy} to D_{uv}, but the Change of Variables Theorem requires a mapping which maps D_{uv} to D_{xy}. In particular, we actually require the inverse $G = F^{-1}$ of our mapping. Solving for x and y we find that

$$(x,y) = G(u,v) = F^{-1}(u,v) = (u+2v, 3u+v)$$

Hence, $\dfrac{\partial(x,y)}{\partial(u,v)} = -5$, and the integrand becomes $x+y = 4u+3v$. Then, the Change of Variables Theorem gives

$$\iint_{D_{xy}} (x+y)\, dx\, dy = \iint_{D_{uv}} (4u+3v)\, |-5|\, du\, dv$$

It is straightforward to write this double integral as an iterated integral and evaluate it. The final result is

$$\iint_{D_{xy}} (x+y)\, dA = \frac{35}{2}$$

EXERCISE 1 Fill in the details in Example 1.

EXAMPLE 2 Use the mapping $(u, v) = F(x, y) = (x + y, -x + y)$ to evaluate

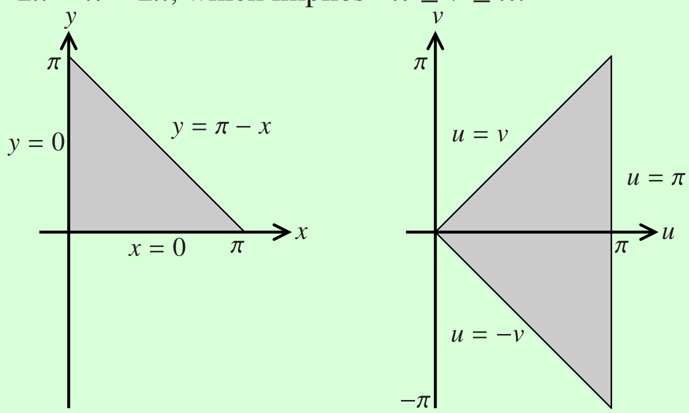
$$\int_0^{\pi} \int_0^{\pi - y} (x + y) \cos(x - y) \, dx \, dy$$

Solution: The region D_{xy} of integration is $0 \leq x \leq \pi - y$ and $0 \leq y \leq \pi$. Thus, the region is bounded by the lines $x = 0$, $y = 0$ and $x = \pi - y$. Since we want to integrate with respect to u and v, we use the mapping F to determine the new region D_{uv}.

For the line $x = 0$, $0 \leq y \leq \pi$, we get $v = y = u$ with $0 \leq u \leq \pi$.

For the line $y = 0$, $0 \leq x \leq \pi$, we get $v = -x = -u$ with $0 \leq u \leq \pi$.

For the line $x + y = \pi$, $0 \leq x \leq \pi$, we get $u = \pi$. We also have $u - v = 2x$. Thus, $v = u - 2x = \pi - 2x$, which implies $-\pi \leq v \leq \pi$.



To apply the Change of Variables Theorem, we also require the Jacobian $\dfrac{\partial(x, y)}{\partial(u, v)}$. Rather than finding the inverse mapping, we can instead use the inverse property of the Jacobian. We have

$$\frac{\partial(u, v)}{\partial(x, y)} = \begin{vmatrix} 1 & 1 \\ -1 & 1 \end{vmatrix} = 2$$

Thus,

$$\frac{\partial(x, y)}{\partial(u, v)} = \frac{1}{\dfrac{\partial(u, v)}{\partial(x, y)}} = \frac{1}{2}$$

Since Jacobian is non-zero and the mapping has continuous partial derivatives, we can apply the Change of Variables Theorem to get

$$\int_0^{\pi} \int_0^{\pi - y} (x + y) \cos(x - y) \, dx \, dy = \iint_{D_{uv}} u \cos(-v) \left| \frac{1}{2} \right| \, dA$$

From the diagram, we observe that we can write the region D_{uv} as $-u \leq v \leq u$ and $0 \leq u \leq \pi$. Thus,

$$\begin{aligned} \int_0^{\pi} \int_0^{\pi - y} (x + y) \cos(x - y) \, dx \, dy &= \int_0^{\pi} \int_{-u}^{u} u \cos(-v) \left| \frac{1}{2} \right| \, dv \, du \\ &= \frac{1}{2} \int_0^{\pi} -u \sin(-v) \Big|_{-u}^{u} \, du \\ &= \frac{1}{2} \int_0^{\pi} 2u \sin u \, du = -u \cos u + \sin u \Big|_0^{\pi} = \pi \end{aligned}$$

Double Integrals in Polar Coordinates

[Refer to Appendix B for an introduction to polar coordinates.]

If the boundary of the region is a circle centered on the origin or a circle that passes through the origin, it will often help to transform from polar to Cartesian coordinates. Recall that the mapping from polar to Cartesian coordinates is

$$(x, y) = F(r, \theta) = (r\cos\theta, r\sin\theta)$$

which has Jacobian,

$$\frac{\partial(x, y)}{\partial(r, \theta)} = r$$

Hence, we must restrict $r > 0$ so that the mapping is one-to-one and the Jacobian is non-zero so that we can apply the Change of Variables Theorem. Note that we can make this restriction even if the origin is in the region as the integral over a single point is 0.

The Change of Variables Theorem in polar coordinates reads:

$$\iint\limits_{D_{xy}} f(x, y)\, dA = \iint\limits_{D_{r\theta}} f(r\cos\theta, r\sin\theta) r\, dr\, d\theta$$

EXAMPLE 3 Evaluate $\displaystyle\iint\limits_{D_{xy}} \frac{x}{x^2 + y^2}\, dA$ where D_{xy} is the half disc $(x-1)^2 + y^2 \leq 1$, $x \geq 1$.

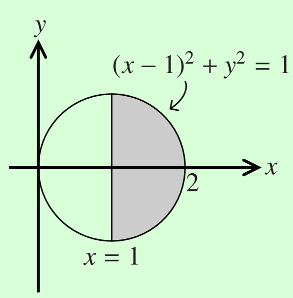
Solution: We first convert the equations from Cartesian coordinates to polar coordinates. Since $x = r\cos\theta$ we get that $x = 1$ becomes

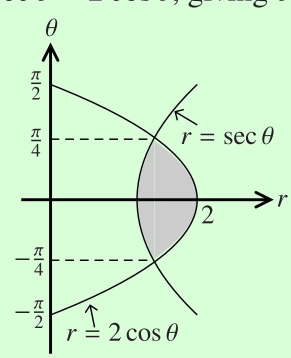
$$\begin{aligned} r\cos\theta &= 1 \\ r &= \sec\theta \end{aligned}$$

Similarly, $x^2 + y^2 = 2x$ becomes

$$\begin{aligned} r^2 &= 2r\cos\theta \\ r &= 2\cos\theta \end{aligned}$$

assuming $r \neq 0$. The image $D_{r\theta}$ is shown in the figure below. The values of θ at the points of intersection are obtained by solving $\sec\theta = 2\cos\theta$, giving $\theta = \pm\dfrac{\pi}{4}$.





The Change of Variables Theorem thus implies

$$\iint\limits_{D_{xy}} \frac{x}{x^2+y^2}\,dx\,dy = \iint\limits_{D_{r\theta}} \frac{r\cos\theta}{r^2}|r|\,dr\,d\theta = \iint\limits_{D_{r\theta}} \cos\theta\,dr\,d\theta.$$

The set $D_{r\theta}$ is described by the inequalities

$$\sec\theta \le r \le 2\cos\theta, \qquad \text{and} \qquad -\frac{\pi}{4} \le \theta \le \frac{\pi}{4}$$

We can thus write the integral over $D_{r\theta}$ as an iterated integral,

$$\iint\limits_{D_{r\theta}} \cos\theta\,dr\,d\theta = \int_{-\frac{\pi}{4}}^{\frac{\pi}{4}} \int_{\sec\theta}^{2\cos\theta} \cos\theta\,dr\,d\theta$$

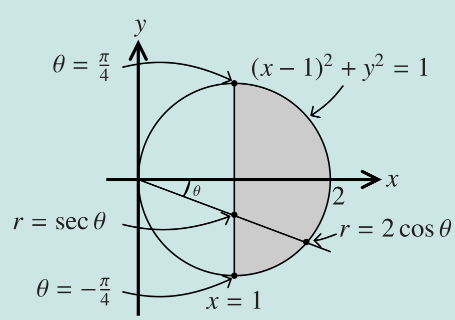
It is a routine matter to evaluate this, leading to the final answer

$$\iint\limits_{D_{xy}} \frac{x}{x^2+y^2}\,dx\,dy = 1$$

EXERCISE 2 Fill in the details in Example 3.

REMARK

Because polar coordinates have a simple geometric interpretation one can obtain the r and θ limits of integration directly from the diagram in the xy-plane, without drawing the region $D_{r\theta}$. The method is illustrated in the diagram.



EXERCISE 3 Evaluate

$$\iint\limits_{D_{xy}} \frac{1}{\sqrt{x^2+y^2}}\,dA$$

where D is the region in the first quadrant bounded by the circles $x^2+y^2 = 1$ and $x^2+y^2 = 4$. Use polar coordinates, as in Example 3.

EXERCISE 4 Evaluate

$$I = \iint_{D_{xy}} xy \, dA$$

where D_{xy} is the set in the first quadrant bounded by $y = x$, $y = ex$, $xy = 2$ and $xy = 3$.

Hint: Find a mapping which maps D_{xy} into a rectangle D_{uv} in the uv-plane.

Chapter 13 Problem Set

1. Find the Jacobian of the mapping

$$(u, v) = F(x, y) = (x^2 \sin y, y^2 \cos x)$$

2. Consider the map defined by

$$(u, v) = F(x, y) = (y + e^{-x}, y - e^{-x})$$

   (a) Show that F has an inverse map by finding F^{-1} explicitly.

   (b) Find the derivative matrices $DF(x, y)$ and $DF^{-1}(u, v)$ and verify that

$$DF(x, y)DF^{-1}(u, v) = I$$

   (c) Verify that the Jacobians satisfy

$$\frac{\partial(x, y)}{\partial(u, v)} = \left[\frac{\partial(u, v)}{\partial(x, y)}\right]^{-1}$$

3. Consider the map defined by

$$(u, v) = F(x, y) = (y + xy, y - xy)$$

   (a) Show that F has an inverse map by finding F^{-1} explicitly.

   (b) Find the derivative matrices $DF(x, y)$ and $DF^{-1}(u, v)$ and verify that

$$DF(x, y)DF^{-1}(u, v) = I$$

   (c) Verify that the Jacobians satisfy

$$\frac{\partial(x, y)}{\partial(u, v)} = \left[\frac{\partial(u, v)}{\partial(x, y)}\right]^{-1}$$

4. Calculate the Jacobian $\dfrac{\partial(u, v)}{\partial(x, y)}$ for the following maps T. Find all points at which the Jacobian is zero. Use the Inverse Map Theorem to prove that T^{-1} exists in a neighbourhood of the indicated point:

   (a) $(u, v) = T(x, y) = (\cos(x+y), \sin(x-y))$; $\left(\frac{\pi}{4}, \frac{\pi}{4}\right)$

   (b) $(u, v) = T(x, y) = (x + y, 2xy^2)$; $(0, 1)$

5. Calculate the approximate area of the image of a small rectangle of area $\Delta x \Delta y$ located at the point (a, b) under the map T defined by

   (a) $T(x, y) = (xy, x^2 - y^2)$, $(a, b) = \left(1, \frac{1}{2}\right)$

   (b) $T(x, y) = \left(\sqrt{x^2 + y^2}, \dfrac{x}{\sqrt{x^2 + y^2}}\right)$, $(a, b) = (1, 1)$

6. Invent a transformation that maps the parallelogram bounded by the lines $y = 3x - 4$, $y = 3x$, $y = \frac{1}{2}x$ and $y = \frac{1}{2}(x + 4)$ onto the unit square in the first quadrant.

7. Invent a transformation that maps the ellipse $x^2 + 4xy + 5y^2 = 4$ onto the unit circle.

8. Invent an invertible transformation that transforms the ellipse $x^2 + 4xy + 5y^2 = 5$ onto the unit circle and determine the inverse map.

9. Invent an invertible transformation that transforms the ellipse $3x^2 + 6xy + 4y^2 = 4$ onto the unit circle and determine the inverse map.

10. Invent a transformation that maps the ellipsoid $x^2 + 8y^2 + 6z^2 + 4xy - 2xz + 4yz = 9$ onto the unit sphere.

11. Invent an invertible transformation that maps the ellipsoid $x^2 + 2y^2 + 2z^2 + 2xy + 2xz + 2yz = 1$ onto the unit sphere.

12. Consider the map defined by

$$(u, v) = F(x, y) = (x + ky^2, y)$$

    where k is a non-negative constant.

    (a) Find the image of the family of lines $x =$ constant under the map. Illustrate with a sketch. What happens when k is close to zero, and when k is very large? Estimate the area of the image of a small rectangle of area $\Delta x \Delta y$.

    (b) Find and sketch the image of the disc $x^2 + y^2 \leq 1$ under the map. How does the value of k affect the image? Make a conjecture about the area of the image.

13. Evaluate the following integrals.

(a) $\iint_D x + y^2 \, dA$ where D is the region bounded by $x = y$, $x = -y$ and $y = -1$.

(b) $\iint_{D_{xy}} (x + 2y)^2 \, dA$, where D_{xy} is bounded by the ellipse $x^2 + 4xy + 5y^2 = 5$.

14. Use the map $T(x, y) = (x + y, -x + y)$ to evaluate

$$\int_0^{\pi} \int_0^{\pi-y} (x + y)\cos(x - y) \, dx \, dy$$

15. Let D be the unit disc $x^2 + y^2 \leq 1$. Use polar coordinates to show that

$$\iint_D e^{x^2+y^2} dA = \pi(e - 1)$$

16. Evaluate $\iint_D \frac{x}{\sqrt{x^2 + y^2}} \, dA$, where D is the region inside the circle $x^2 + y^2 = 2x$, but outside the circle $x^2 + y^2 = 1$. Use polar coordinates and describe the image of D.

17. Let D be the region in the xy-plane enclosed by the lines $y = 2-x$, $y = 4-x$, $y = x$ and $y = 0$. Evaluate the Jacobian of the map $(x, y) = F(u, v) = (u + uv, u - uv)$, and show that it is never zero on D. Sketch the image of D in the uv-plane. Use this map to evaluate

$$\iint_D \frac{e^{\frac{x-y}{x+y}}}{x + y} \, dx \, dy$$

18. Let

$$f(x, y) = \begin{cases} 1 & \text{if } x + y \geq 1 \\ -1 & \text{if } x + y < 1 \end{cases}$$

Evaluate $\iint_D f(x, y) \, dA$, where D is the subset of $\mathbb{R}^2$ defined by $|x| + |y| \leq 2$.

19. A metal plate, bounded by $x^2 - y^2 = 1$, $-x^2 + 3y^2 = 1$, $x = 0$ and $y = 0$, and lying in the first quadrant, is coated with silver. The density of silver at position (x, y) on the plate is given by $\rho(x, y) = xy$ grams per unit area. Calculate the total amount of silver on the plate.

20. Let D_{xy} be the region bounded by $y = 1 - x$, $y = 2 - x$, $y = 0$ and $x = y$ and let $(u, v) = F(x, y) = \left(x - y, \frac{1}{x+y}\right)$.

(a) Sketch the image of D under F in the uv-plane.

(b) Find the Jacobian of F and show that it is never 0 on D.

(c) Find the mapping F^{-1} and the Jacobian for F^{-1}.

(d) Use the mapping F to evaluate $\iint_{D_{xy}} \frac{x - y}{x + y} \, dA$.

21. Find a linear transformation that maps the ellipse $x^2 + 4xy + 5y^2 = 4$ onto a unit circle. Hence show that the area enclosed by the ellipse equals 4π square units (without explicitly integrating).

22. Let D be the subset of $\mathbb{R}^2$ defined by $|x| + |y| \leq 1$, and let f be a continuous single-variable function on the interval $[-1, 1]$. Prove that

$$\iint_D f(x + y) \, dx \, dy = \int_{-1}^{1} f(u) \, du$$

23. Let D be the disc of radius b centered at the origin, and let f be a continuous single-variable function. Prove that

$$\iint_D f(x^2 + y^2) \, dA = \pi \int_0^{b^2} f(u) \, du$$

24. Consider the regions $D_{xy} = \{(x, y) \mid x^2+4xy+13y^2 \leq 9\}$ and $D_{uv} = \{(u, v) \mid u^2 + v^2 \leq 1\}$.

(a) Find an invertible mapping F that transforms D_{xy} into D_{uv}. Prove that your mapping F is invertible.

(b) The number of bacteria per unit area in D_{xy} is given by

$$c(x, y) = \frac{10}{9\pi}(x^2 + 4xy + 13y^2)^2$$

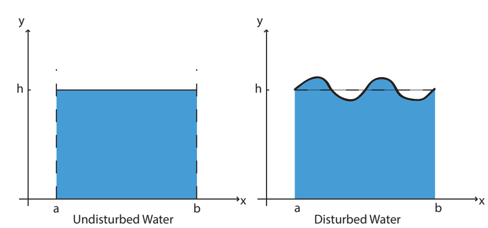
Use the Change of Variables Theorem to write an expression for the number of bacteria in D_{xy} as double integral over D_{uv}.

(c) Determine the number of bacteria in D_{xy}.

25. * Let f be a continuous single-variable function. Prove that

$$2\int_a^b \int_a^x f(x)f(y) \, dy \, dx = \left[\int_a^b f(x) \, dx\right]^2$$

26. Consider a large tank with water of depth h. See the left-hand figure — we'll ignore the third dimension here.



Suppose now that the water is disturbed, creating the wave shown in the right-hand figure, and that the surface of the water has the equation

$$y = h + A \sin k(x - a), \quad k = 4\pi/(b - a)$$

(a) The *average depth* of the water over the interval $a \le x \le b$ is given by $\dfrac{\int_a^b f(x)dx}{b-a}$ where $y = f(x)$ defines the surface of the water. Calculate the average depth for the disturbed and undisturbed cases, and compare your answers. Does your answer agree with your physical intuition? *Suggestion:* Without loss of generality, set $a = 0$.

(b) The *centre of mass* of the water is located at the point $(\bar{x}, \bar{y})$ where

$$\bar{x} = \frac{\iint_D \rho x \, dxdy}{\iint_D \rho \, dxdy}, \quad \bar{y} = \frac{\iint_D \rho y \, dxdy}{\iint_D \rho \, dxdy}$$

where ρ is the density of water (assumed constant) and D is the region occupied by the water. Calculate the centre of mass for the disturbed and undisturbed cases, and compare your answers. You may be surprised by the result!

27. * (a) Show that $\displaystyle\iint_{D(R)} e^{-(x^2+y^2)} \, dx\, dy = \pi(1 - e^{-R^2})$ where $D(R)$ is the disc of radius R, centre $(0,0)$.

(b) Let D be the square $\{(x, y) \mid |x| \le b, |y| \le b\}$. Show that

$$\iint_D e^{-(x^2+y^2)} \, dx\, dy = 4\left(\int_0^b e^{-x^2} \, dx\right)^2 .$$

(c) Hence, prove that

$$\int_0^\infty e^{-x^2} dx = \frac{\sqrt{\pi}}{2}$$

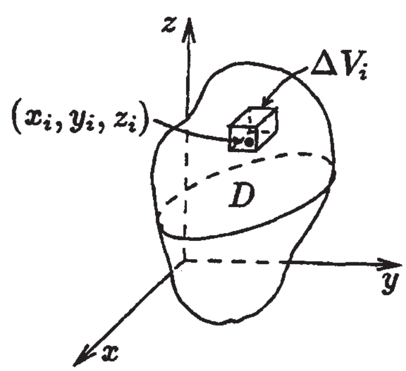
a result which is important in probability theory and in other applications. This integral cannot be evaluated directly.

Chapter 14

Triple Integrals

14.1 Definition of Triple Integrals

Let D be a closed bounded set in $\mathbb{R}^3$ whose boundary consists of a finite number of surface elements which are smooth except possibly at isolated points. Let $f(x, y, z)$ be a function which is bounded on D. Subdivide D by means of three families of planes which are parallel to the $xy-$, $yz-$, and $xz-$planes respectively, forming a partition P of D.



Label the N rectangular blocks that lie completely in D in some specific order, and denote their volumes by $\Delta V_i, i = 1, \ldots, n$. Choose an arbitrary point (x_i, y_i, z_i) in the i-th block, $i = 1, \ldots, n$, and form the Riemann sum

$$\sum_{i=1}^{n} f(x_i, y_i, z_i)\Delta V_i \tag{14.1}$$

Let ΔP denote the maximum of the dimensions of all rectangular blocks in the partition P.

DEFINITION
Integrable

A function $f(x, y, z)$ which is bounded on a closed bounded set $D \subset \mathbb{R}^3$ is said to be **integrable** on D if and only if all Riemann sums approach the same value as $\Delta P \to 0$.

DEFINITION
Triple Integral

If $f(x, y, z)$ is integrable on a closed bounded set D, then we define the **triple integral** of f over D, as

$$\iiint_D f(x, y, z)\, dV = \lim_{\Delta P \to 0} \sum_{i=1}^{n} f(x_i, y_i, z_i)\Delta V_i$$

Is there any guarantee that the limiting process in the definition of the triple integral actually leads to a unique value, i.e. that the limit exists? It is possible to define weird functions for which the limit does not exist, i.e. which are not integrable on D. However, if f is continuous on D, it can be proved that f is integrable on D. Functions which are discontinuous in D may be integrable on D. For example, if f is continuous on D except at points which lie on a surface or curve in D, then f is integrable on D. The proofs of these results are beyond the scope of this course.

Interpretation of the Triple Integral

When you encounter the triple integral symbol

$$\iiint_D f(x, y, z)\, dV$$

you should think of "limit of a sum". In itself, the triple integral is a mathematically defined object. It has many interpretations, depending on the interpretation that you assign to the integrand $f(x, y, z)$. The "dV" in the triple integral symbol should remind you of the volume of a rectangular block in a partition of D.

Triple Integral as Volume:

The simplest interpretation is when you specialize f to be the constant function with value unity:

$$f(x, y, z) = 1, \quad \text{for all } (x, y, z) \in D$$

Then, the Riemann sum (14.1) simply sums the volumes of all rectangular blocks in D, and the triple integral over D serves to define the volume $V(D)$ of the set D:

$$V(D) = \iiint_D 1\, dV$$

Triple Integral as Mass:

Think of a planet or star whose density varies with position. Let D denote the subset of $\mathbb{R}^3$ occupied by the star. Let $f(x, y, z)$ denote the density (mass per unit volume) at position (x, y, z). The mass of a small rectangular block located within the star at position (x_i, y_i, z_i) will be approximately

$$\Delta M_i \approx f(x_i, y_i, z_i)\Delta V_i$$

Thus, the Riemann sum corresponding to a partition P of D

$$\sum_{i=1}^{n} f(x_i, y_i, z_i)\Delta V_i$$

will approximate the total mass M of the star, and the triple integral of f over D, being the limit of the Riemann sum, will represent the total mass:

$$M = \iiint_D f(x, y, z)\, dV$$

Average Value of a Function:

By analogy with functions of one and two variables we can use the triple integral to define the average value of a function $f(x, y, z)$ over a closed and bounded set $D \subset \mathbb{R}^3$.

DEFINITION
Average Value

Let $D \subset \mathbb{R}^3$ be closed and bounded with volume $V(D) \neq 0$, and let $f(x, y, z)$ be a bounded and integrable function on D. The **average value** of f over D is defined by

$$f_{avg} = \frac{1}{V(D)} \iiint_D f(x, y, z)\, dV$$

REMARK

If you have the impression that you have read this section someplace else, you're right. Compare it with Section 11.1. The only essential change is to replace "area" by "volume".

Properties of the Triple Integral

Of course, the triple integral satisfies the same basic properties as the double integral.

THEOREM 1 **(Linearity)**

If $D \subset \mathbb{R}^3$ is a closed and bounded set, c is a constant, and f and g are two integrable functions on D, then

$$\iiint_D (f + g)\, dV = \iiint_D f\, dV + \iiint_D g\, dV$$

$$\iiint_D cf\, dV = c \iiint_D f\, dV$$

THEOREM 2 **(Basic Inequality)**

If $D \subset \mathbb{R}^3$ is a closed and bounded set and f and g are two integrable functions on D such that $f(x, y, z) \leq g(x, y, z)$ for all $(x, y, z) \in D$, then

$$\iiint_D f\, dV \leq \iiint_D g\, dV$$

THEOREM 3 **(Absolute Value Inequality)**

If $D \subset \mathbb{R}^3$ is a closed and bounded set and f is an integrable function on D, then

$$\left| \iiint_D f \, dV \right| \leq \iiint_D |f| \, dV$$

THEOREM 4 **(Decomposition)**

Assume $D \subset \mathbb{R}^3$ is a closed and bounded set and f is an integrable function on D. If D is decomposed into two closed and bounded subsets D_1 and D_2 by a piecewise smooth surface C, then

$$\iiint_D f \, dV = \iiint_{D_1} f \, dV + \iiint_{D_2} f \, dV$$

14.2 Iterated Integrals

We generalize the method used in Section 11.2, and show how to express a triple integral as a 3-fold iterated integral. This enables you to evaluate triple integrals exactly for sufficiently simple functions and integration sets.

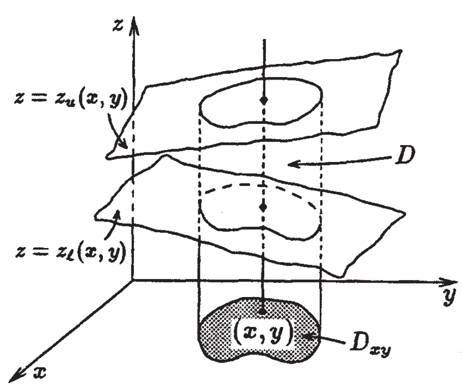
Consider a set $D \subset \mathbb{R}^3$ which is described by inequalities of the form

$$z_\ell(x, y) \leq z \leq z_u(x, y)$$

and

$$(x, y) \in D_{xy}$$

Here D_{xy} is a closed bounded subset of $\mathbb{R}^2$ whose boundary is a piecewise smooth closed curve, and z_ℓ, z_u are continuous functions on D_{xy}. Think of the set D as being the 3-D region with bottom surface $z = z_\ell(x, y)$ and top surface $z = z_u(x, y)$, where the extent is defined by the 2-D set D_{xy}.



In order to write a triple integral as an iterated integral, take an arbitrary point $(x, y) \in D_{xy}$. Then you integrate $f(x, y, z)$ with respect to z from $z_\ell(x, y)$ to $z_u(x, y)$, and integrate the result over D_{xy}, as a double integral.

This procedure essentially sums over all rectangular blocks in a partition of D, and hence gives the triple integral of $f(x, y, z)$ over D.

THEOREM 1 Let D be the subset of $\mathbb{R}^3$ defined by

$$z_\ell(x,y) \le z \le z_u(x,y) \quad \text{and} \quad (x,y) \in D_{xy}$$

where z_ℓ and z_u are continuous functions on D_{xy}, and D_{xy} is a closed bounded subset in $\mathbb{R}^2$, whose boundary is a piecewise smooth closed curve. If $f(x,y,z)$ is continuous on D, then

$$\iiint\limits_D f(x,y,z)\,dV = \iint\limits_{D_{xy}} \int\limits_{z_\ell(x,y)}^{z_u(x,y)} f(x,y,z)\,dz\,dA$$

REMARK

As with double iterated integrals, we are doing partial integration. That is, to evaluate the inner integral of

$$\iint\limits_{D_{xy}} \int\limits_{z_\ell(x,y)}^{z_u(x,y)} f(x,y,z)\,dz\,dA$$

we hold x and y constant and integrate with respect to z.

Keep in mind that when evaluating a triple integral, it is not essential to integrate first with respect to z. One chooses the order of integration that is most convenient.

That is, if you can describe D by inequalities of the form

$$x_\ell(y,z) \le x \le x_u(y,z)$$

with $(y,z) \in D_{yz}$, then you would get

$$\iiint\limits_D f(x,y,z)\,dV = \iint\limits_{D_{yz}} \int\limits_{x_\ell(y,z)}^{x_u(y,z)} f(x,y,z)\,dx\,dA$$

On the other hand, if you can describe D by inequalities of the form

$$y_\ell(x,z) \le y \le y_u(x,z)$$

with $(x,z) \in D_{xz}$, then you would get

$$\iiint\limits_D f(x,y,z)\,dV = \iint\limits_{D_{xz}} \int\limits_{y_\ell(x,z)}^{y_u(x,z)} f(x,y,z)\,dy\,dA$$

We will demonstrate this in the example and exercises below.

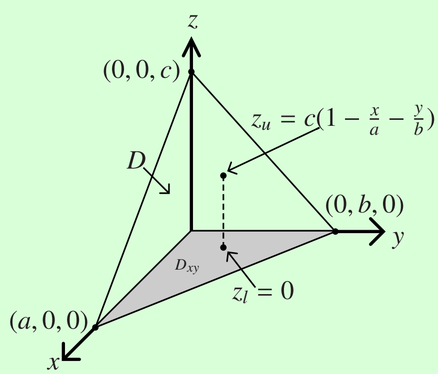
EXAMPLE 1 Evaluate $\displaystyle\iiint\limits_D z\,dV$, where D is the solid tetrahedron with vertices $(a,0,0)$, $(0,b,0)$, $(0,0,c)$, and $(0,0,0)$.

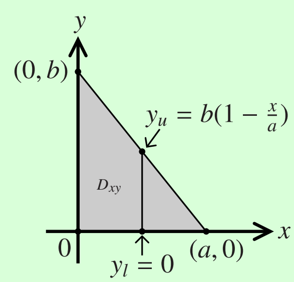
Solution: The tetrahedron is bounded by the planes $x = 0$, $y = 0$, $z = 0$, and $\dfrac{x}{a} + \dfrac{y}{b} + \dfrac{z}{c} = 1$. Thus, the region D can be described by

$$0 \le z \le c\left(1 - \frac{x}{a} - \frac{y}{b}\right), \qquad \text{and} \qquad (x,y) \in D_{xy}$$

where D_{xy} is bounded by $x = 0$, $y = 0$, and the intersection of the inclined face with $z = 0$. The intersection is

$$\frac{x}{a} + \frac{y}{b} + \frac{0}{c} = 1 \Rightarrow y = b\left(1 - \frac{x}{a}\right)$$





Thus, by Theorem 1,

$$\iiint\limits_D z\,dV = \iint\limits_{D_{xy}} \int_0^{c\left(1-\frac{x}{a}-\frac{y}{b}\right)} z\,dz\,dA = \int_0^a \int_0^{b\left(1-\frac{x}{a}\right)} \int_0^{c\left(1-\frac{x}{a}-\frac{y}{b}\right)} z\,dz\,dy\,dx$$

on writing the outer double integral over D_{xy} as a double iterated integral. After evaluating the integrals, one obtains as a final answer,

$$\iiint_D z\,dV = \frac{1}{24}abc^2$$

EXERCISE 1 Verify the answer in Example 1 by evaluating the iterated integral.

EXERCISE 2 Write the triple integral in Example 1 as an iterated integral taking the variables in the order y, x, z. Evaluate the iterated integral and verify you get the same answer as in Example 1.

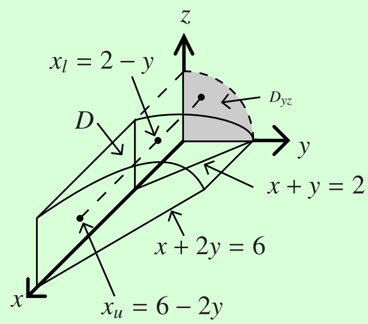
EXERCISE 3 In how many ways can a triple integral be written as an iterated integral?

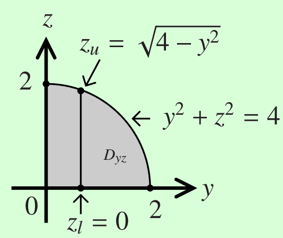
EXAMPLE 2 Evaluate $\displaystyle\iiint_D \frac{z}{4-y}\,dV$, where D is the region bounded by the cylinder $y^2+z^2=4$, and the planes $x+y=2$, $x+2y=6$, $z=0$, $y=0$, and lying in the first octant.

Solution: Since x only occurs in two of the equations, it is convenient to integrate first with respect to x, and describe D by the inequalities

$$2-y \le x \le 6-2y \quad \text{and} \quad (y,z)\in D_{yz}$$

where D_{yz} is the region in the first quadrant bounded by $y^2+z^2=4$, $y=0$, and $z=0$.





Thus,

$$\iiint_D \frac{z}{4-y}\,dV = \iint_{D_{yz}} \int_{2-y}^{6-2y} \frac{z}{4-y}\,dx\,dA = \iint_{D_{yz}} z\,dA,$$

$$= \int_0^2 \int_0^{\sqrt{4-y^2}} z\,dz\,dy = \frac{1}{2}\int_0^2 (4-y^2)\,dy = \frac{8}{3}$$

EXERCISE 4 Evaluate the triple integral in Example 2 by writing it as an iterated integral with the variables in the order z, x, y. Why would it not make sense to integrate first with respect to y?

EXERCISE 5 Let D be the subset of $\mathbb{R}^3$ (a prism) bounded by the planes $x=0$, $x=2$, $y=0$, $z=0$, and $y+z=1$. Evaluate $\displaystyle\iiint_D y\,dV$.

14.3 The Change of Variables Theorem for Triple Integrals

A mapping F from $\mathbb{R}^3$ to $\mathbb{R}^3$ can be used to simplify a triple integral

$$\iiint_{D_{xyz}} f(x,y,z)\,dV$$

either by changing the integrand $f(x,y,z)$ or by deforming the set D_{xyz} in xyz-space into a simpler shape D_{uvw} in uvw-space, thereby simplifying the limits of integration. In this type of calculation, it is convenient to replace the symbol "dV" in the triple integral by "$dx\,dy\,dz$" if one is working in xyz-space, and by "$du\,dv\,dw$" if one is working in uvw-space.

Just as for change of variables in double integrals, we require an appropriate Jacobian in this context. The details are completely analogous to the two-variable case and have been placed in an appendix to this section. With this in hand, we can now state the Change of Variables Theorem for triple integrals.

THEOREM 1 **(Change of Variables Theorem)**

Let

$$x = g(u,v,w), \quad y = h(u,v,w), \quad z = k(u,v,w)$$

be a one-to-one mapping of D_{uvw} onto D_{xyz}, with g, h, k having continuous partials, and

$$\frac{\partial(x,y,z)}{\partial(u,v,w)} \neq 0 \quad \text{on} \quad D_{uvw}$$

If $f(x,y,z)$ is continuous on D_{xyz}, then

$$\iiint_{D_{xyz}} f(x,y,z)\,dV = \iiint_{D_{uvw}} f\Big(g(u,v,w), h(u,v,w), k(u,v,w)\Big)\left|\frac{\partial(x,y,z)}{\partial(u,v,w)}\right|\,dV$$

A proof is beyond the scope of this course, but the volume transformation formula using the Jacobian in the appendix to this section makes the theorem plausible, as in the case of the double integral.

EXAMPLE 1 Evaluate $I = \iiint_{D_{xyz}} x^2\,dV$, where D_{xyz} is the subset of $\mathbb{R}^3$ bounded by the surfaces $xy = 1$, $xy = 3$, and the planes $y + z = -1$, $y + z = 0$, $x + y + z = 1$ and $x + y + z = 2$.

Solution: This solid is difficult to draw, but one can visualize it, since it is bounded by level surfaces of three functions, namely

$$xy, \quad y + z, \quad \text{and} \quad x + y + z$$

Thus, the solid D_{xyz} is described by the inequalities

$$1 \le xy \le 3, \quad -1 \le y+z \le 0, \quad 1 \le x+y+z \le 2 \tag{14.2}$$

This suggests that we define a mapping

$$u = xy, \quad v = y+z, \quad w = x+y+z \tag{14.3}$$

The Jacobian is

$$\frac{\partial(u,v,w)}{\partial(x,y,z)} = \det\begin{bmatrix} y & x & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = x$$

By the Change of Variables Theorem,

$$I = \iiint\limits_{D_{xyz}} x^2\, dx\, dy\, dz = \iiint\limits_{D_{uvw}} x^2 \left|\frac{\partial(x,y,z)}{\partial(u,v,w)}\right| du\, dv\, dw$$

By the inverse property of the Jacobian,

$$\frac{\partial(x,y,z)}{\partial(u,v,w)} = \left[\frac{\partial(u,v,w,)}{\partial(x,y,z)}\right]^{-1} = \frac{1}{x}$$

It follows from the inequalities (14.2) that $x > 0$ on D_{xyz}. Thus, equation (14.3) gives

$$I = \iiint\limits_{D_{uvw}} x\, du\, dv\, dw$$

The next step is to express the integrand x in terms of u, v, w. It follows from equations (14.3) that $x = w - v$. Hence,

$$I = \int\limits_{D_{uvw}} (w - v)\, du\, dv\, dw \tag{14.4}$$

The inequalities (14.2) imply that the image of the set D_{xyz} under the mapping (14.3) is the rectangular block D_{uvw} defined by

$$1 \le u \le 3, \quad -1 \le v \le 0, \quad 1 \le w \le 2$$

Therefore, we can write the triple integral (14.4) as an iterated integral, and since D_{uvw} is rectangular, the order is immaterial:

$$I = \int_1^2 \int_{-1}^0 \int_1^3 (w - v)\, du\, dv\, dw = \cdots = 4$$

EXERCISE 1 Verify the result $\dfrac{\partial(u,v,w)}{\partial(x,y,z)} = x$ in Example 1.

EXERCISE 2 Find the volume of the solid bounded by the six planes $x+y=1$, $x+y=2$, $x-y=-1$, $x-y=1$, $x+y+z=0$, $x+y+z=3$.

In double integrals we saw that if there is symmetry about the origin it may be helpful to evaluate the double integral using polar coordinates. Similarly, if we have symmetry about the z-axis or the origin in $\mathbb{R}^3$ it may be helpful to use our mappings to *cylindrical coordinates* or *spherical coordinates*. Refer to Appendix B for an introduction to these coordinate systems.

Triple Integrals in Cylindrical Coordinates

Recall that the mapping from Cartesian coordinates to cylindrical coordinates is

$$x = r\cos\theta, \quad y = r\sin\theta, \quad z = z$$

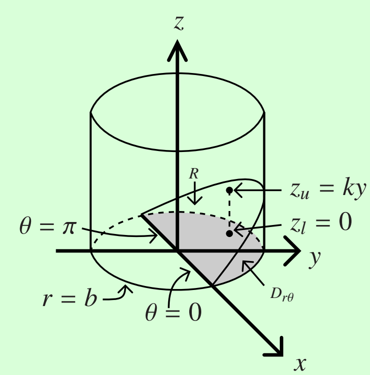
with $r \geq 0$, $0 \leq \theta < 2\pi$, and the Jacobian is $\dfrac{\partial(x,y,z)}{\partial(r,\theta,z)} = r$ (verify). Since we need $\dfrac{\partial(x,y,z)}{\partial(r,\theta,z)} \neq 0$, we must again restrict $r > 0$. So for cylindrical coordinates, the formula in the Change of Variables Theorem reads

$$\iiint_{D_{xyz}} f(x,y,z)\,dx\,dy\,dz = \iiint_{D_{r\theta z}} f(r\cos\theta, r\sin\theta, z)r\,dr\,d\theta\,dz$$

EXAMPLE 2 A wedge is cut from the cylinder $x^2+y^2=b^2$, by the planes $z=0$ and $z=ky$, where b and k are positive constants, and y is assumed to be non-negative. Find the volume of the wedge.

Solution: The volume V is given by

$$V = \iiint_R 1\,dV$$



In cylindrical coordinates we have the cylinder $r=b$, the plane $z=0$ and the plane $z = kr\sin\theta$. Hence, the solid is described by

$$0 \leq z \leq kr\sin\theta, \quad 0 \leq r \leq b, \quad 0 \leq \theta \leq \pi$$

Using the Change of Variables Theorem gives

$$V = \int_0^{\pi}\int_0^{b}\int_0^{kr\sin\theta} r\,dz\,dr\,d\theta = \int_0^{\pi}\int_0^{b} kr^2\sin\theta\,dr\,d\theta$$
$$= \int_0^{\pi} \frac{1}{3}kb^3\sin\theta\,d\theta$$
$$= \frac{2}{3}kb^3$$

EXERCISE 3

The density μ of the contents of a cylindrical drum defined by

$$x^2 + y^2 \leq 1 \quad \text{and} \quad 0 \leq z \leq 2$$

is given by

$$\mu = \frac{k(2-z)}{1 + x^2 + y^2}$$

where k is constant. Find the total mass.

EXERCISE 4

Calculate the volume of the solid enclosed by the paraboloid $z = x^2 + y^2$ and the lower part of the cone $(z-2)^2 = x^2 + y^2$.

Triple Integrals in Spherical Coordinates

Recall that the mapping from spherical coordinates to Cartesian coordinates are

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

with $\rho \geq 0$, $0 \leq \phi \leq \pi$, $0 \leq \theta < 2\pi$. The Jacobian is

$$\frac{\partial(x,y,z)}{\partial(\rho,\theta,\phi)} = \rho^2\sin\phi$$

EXERCISE 5

Verify that $\dfrac{\partial(x,y,z)}{\partial(\rho,\theta,\phi)} = \rho^2\sin\phi$.

Thus, for spherical coordinates, we must restrict $\rho > 0$ and $0 < \phi < \pi$ so that the Jacobian is non-zero and the mapping is one-to-one. Observe that this means we are not just removing one point, but the entire z-axis. However, this still will not affect our result as the triple integral over the z-axis is 0. Hence, the Change of Variables Theorem in spherical coordinates reads:

$$\iiint\limits_{D_{xyz}} f(x,y,z)dV = \iiint\limits_{D_{\rho\theta\phi}} f(\rho\sin\phi\cos\theta, \rho\sin\phi\sin\theta, \rho\cos\phi)\rho^2\sin\phi\; d\rho d\theta d\phi$$

EXAMPLE 3 Evaluate $\iiint\limits_D \frac{1}{x^2+y^2+z^2} dV$ where D is the spherical shell between the spheres of radius a and b centered on the origin ($a < b$).

Solution: You would not succeed in evaluating this triple integral as an iterated integral in terms of x, y, and z. However, if you use spherical coordinates, the calculation is simple. In terms of spherical coordinates ρ, ϕ, θ, the set D is defined by

$$a \le \rho \le b, \quad 0 \le \phi \le \pi, \quad 0 \le \theta \le 2\pi$$

Using the Change of Variables Theorem gives

$$\iiint\limits_D \frac{1}{x^2+y^2+z^2}\, dV = \int_0^{2\pi}\int_0^{\pi}\int_a^b \frac{1}{\rho^2}(\rho^2 \sin\phi)\, d\rho\, d\phi\, d\theta = \cdots = 4\pi(b-a)$$

EXERCISE 6 Calculate the volume of the solid ellipsoid

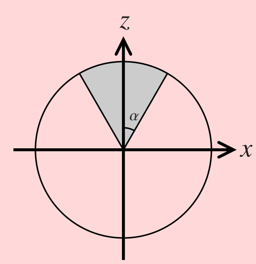
$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \le 1$$

where a, b, c are positive constants.

Hint: Make the change of variables $(x, y, z) = (au, bv, cw)$, and then transform the ellipsoid into a solid sphere.

EXERCISE 7 A conical drill bit, angle α, drills into a solid sphere of radius b until the tip reaches the center. Show that the volume of the solid removed is

$$V(\alpha) = \frac{2}{3}\pi b^3(1 - \cos\alpha)$$



cross section $y = 0$

Appendix: The Jacobian in 3-D

At the end of Section 12.2, we generalized the concept of a mapping F from $\mathbb{R}^2$ to $\mathbb{R}^2$ to a mapping F from $\mathbb{R}^n$ to $\mathbb{R}^m$, and defined the $m \times n$ derivative matrix $DF(\mathbf{x})$. If $m = n$, then we can define the Jacobian of the mapping, as follows.

DEFINITION

Jacobian

For a mapping defined by

$$\mathbf{u} = F(\mathbf{x}) = \big(f_1(\mathbf{x}), \ldots, f_n(\mathbf{x})\big)$$

where $\mathbf{u} = (u_1, \ldots, u_n)$ and $\mathbf{x} = (x_1, \ldots, x_n)$, the **Jacobian** of F is

$$\frac{\partial(u_1, \ldots, u_n)}{\partial(x_1, \ldots, x_n)} = \det[DF(x,y)] = \det \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}$$

We note that the inverse property of the Jacobian also generalizes:

$$\frac{\partial(x_1, \ldots, x_n)}{\partial(u_1, \ldots, u_n)} = \frac{1}{\dfrac{\partial(u_1, \ldots, u_n)}{\partial(x_1, \ldots, x_n)}}$$

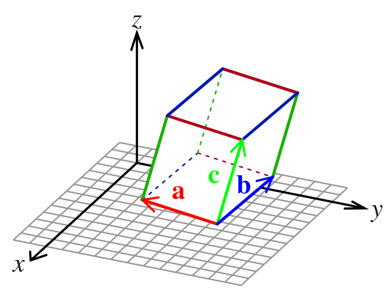
where $\dfrac{\partial(x_1, \ldots, x_n)}{\partial(u_1, \ldots, u_n)}$ is the Jacobian of the inverse mapping of F.

Geometrical Interpretation of the Jacobian in 3-D

The interpretation is based on the following result from linear algebra. The volume of a parallelepiped which is defined by three vectors

$\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$, $\begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$, and $\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$ is given by

$$\text{Volume} = \left| \det \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} \right|$$



Consider a mapping defined by

$$(u, v, w) = F(x, y, z) = \big(f(x,y,z), g(x,y,z), h(x,y,z)\big)$$

The image of a small rectangular block of volume $\Delta V_{xyz} = \Delta x \Delta y \Delta z$ in xyz-space under this mapping can be approximated by a small parallelepiped in uvw-space. As in the 2-D case we can use the linear approximation and the formula above to approximate the volume ΔV_{uvw} of the image. The result is

$$\Delta V_{uvw} \approx \left| \frac{\partial(u,v,w)}{\partial(x,y,z)} \right| \Delta V_{xyz}$$

where $\dfrac{\partial(u,v,w)}{\partial(x,y,z)}$ is the Jacobian of the mapping F evaluated at P.

Chapter 14 Problem Set

1. Write $\iiint\limits_D f(x,y,z)\ dV$ as an iterated integral, for each 3-D region D.
   (a) D is the rectangular box defined by $|x-1| \leq 2, |y| \leq 3, |z+1| \leq 1$.
   (b) D is the cylindrical solid defined by $\frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1, |z-2| \leq 1$.
   (c) D is the tetrahedron with vertices $(a,0,0)$, $(0,b,0)$, $(0,0,c)$ and $(0,0,-c)$.
   (d) D is the "ice-cream cone" bounded by $x^2 + y^2 = \frac{1}{4}z^2$, $z \geq 0$ and the hemisphere defined by $x^2 + y^2 + z^2 = 25$, $z > 0$.
   (e) D is the solid bounded by the paraboloid $y = 1 - x^2 - z^2$, and the hemisphere defined by $x^2 + y^2 + z^2 = 3$, $y < 0$ in D, $y < 1 - x^2 - z^2$.
2. Consider the triple integral $\iiint\limits_D e^x\ dV$, where D is the 3-d region bounded by the planes $x = 0$, $y = 0$, $z = 0$ and $x + y + z = 1$. Write it as an iterated integral in the order z, y, x. Notice that the order x, z, y will give a simpler integration. Evaluate the integral using this order.
3. Evaluate $\iiint\limits_D x^2 + y\ dV$ where D is the region bounded by $x + y + z = 2$, $z = 2$, $x = 1$ and $y = x$.
4. Evaluate $\iiint\limits_{D_{xyz}} e^{x-y+z}\ dV$, where D_{xyz} is bounded by the planes $x - y + z = 2$, $x - y + z = 3$, $x + 2y = -1$, $x + 2y = 1$, $x - z = 0$ and $x - z = 2$.
5. Let $0 < a < b$ and $0 < c < d$. Show that the region D in the first quadrant bounded by $ay = x^3$, $by = x^3$, $cx = y^3$ and $dx = y^3$ has area $\frac{1}{2}(\sqrt{b} - \sqrt{a})(\sqrt{d} - \sqrt{c})$.
6. Evaluate the following integrals.
   (a) $\iiint\limits_D (x^2 + y^2 + z^2)^{-3/2}\ dV$ where D is the bounded by $x^2 + y^2 + z^2 = a^2$ and $x^2 + y^2 + z^2 = b^2$ with $0 < b < a$.
   (b) $\displaystyle\int_0^1 \int_0^{\sqrt{1-x^2}} \int_{\sqrt{x^2+y^2}}^{\sqrt{2-x^2-y^2}} dz\, dy\, dx.$
   (c) $\displaystyle\int_0^1 \int_0^{\sqrt{1-y^2}} \int_{\sqrt{3x^2+3y^2}}^{\sqrt{3}} dz\, dx\, dy.$
7. Find the volume of the region bounded by the surfaces.
   (a) $z = \sqrt{x^2 + y^2}$, $x^2 + y^2 = 4$, $z = 0$.
   (b) $x + y + z = 2$, $x^2 + y^2 = 1$, $z = 0$.
   (c) Inside $x^2 + y^2 + z^2 = 2$ but outside $x^2 + y^2 = 1$.
8. Let D be the region in the first octant bounded by the planes $x = z$, $y = z$, and $z = 1$. Consider $\iiint_D \frac{1}{3}z^3\ dV$.
   (a) Write the triple integral as an iterated integral in the order $dz\, dy\, dx$.
   (b) Write the triple integral as an iterated integral in the order $dx\, dy\, dz$ and evaluate the integral using this order.
9. Describe the 3-d region of integration for the iterated integral
$$\int\limits_{y=0}^{1} \int\limits_{x=y-1}^{1-y} \int_{z=-\sqrt{(1-y)^2-x^2}}^{\sqrt{(1-y)^2-x^2}} f(x,y,z)\, dz\, dx\, dy$$
and find the limits when the order of integration is y, x, z.
10. The temperature at points in the cube
$$C = \{(x,y,z) \mid |x| \leq 1, |y| \leq 1, |z| \leq 1\}$$
is $100r^2$, where r is the distance to the origin. Find the average temperature. At what points of the cube does the temperature equal the average temperature?
11. Determine the volume bounded by the cone $z = 2\sqrt{x^2 + y^2}$ and the paraboloid $z = 1 - 8(x^2 + y^2)$.
12. Use spherical coordinates to evaluate $\iiint\limits_D (x^2 + y^2 + z^2)^{-3/2}\ dV$, where D is the solid bounded by the spheres $x^2 + y^2 + z^2 = a^2$ and $x^2 + y^2 + z^2 = b^2$, with $0 < b < a$.
13. Evaluate the following triple integrals by transforming to spherical coordinates:
   (a) $\displaystyle\int_0^1 \int_0^{\sqrt{1-x^2}} \int_{\sqrt{x^2+y^2}}^{\sqrt{2-x^2-y^2}} dz\, dy\, dx$
   (b) $\displaystyle\int_0^1 \int_0^{\sqrt{1-y^2}} \int_{\sqrt{3x^2+3y^2}}^{\sqrt{3}} dz\, dx\, dy$
14. Calculate the volume enclosed by the cone $z^2 = x^2 + y^2$ and the plane $z = h > 0$, first using cylindrical coordinates, and then using spherical coordinates.
15. Find the mass inside the sphere $x^2 + y^2 + z^2 = 1$, if the density is proportional to (i) the distance from the z-axis (ii) the distance from the xy-plane. Think about both spherical and cylindrical coordinates, and use whichever is simpler.
16. Show that the volume V which lies inside the sphere $x^2 + y^2 + (z-a)^2 = a^2$ and outside the sphere $x^2 + y^2 + z^2 = 4k^2a^2$, where k is a constant, $0 < k < 1$, is given by
$$V = \frac{4\pi}{3}a^3(1 - 4k^3 + 3k^4)$$
17. Let V denote the volume of the first octant region bounded by the coordinate planes and the parabolic cylinders
$$a^2y = b(a^2 - x^2), \quad a^2z = c(a^2 - x^2), \quad a, b, c > 0$$
Show that $V = 8abc/15$.

18. Let V denote the volume of the ellipsoid $\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$. Find V by performing a transformation, but without integrating.

19. Suppose that a hemispherical tank (with the flat part on the ground) with radius R is partially filled with water, so that the depth of the water is h. Find the volume of water in the tank. Consider spherical and cylindrical coordinates before trying to evaluate; one of them will be easier to use.

20. A glacier which occupies the region

$$-\sqrt{10^{-2} - x^2} < z < 0$$

moves parallel to the y-axis with velocity in km/year given by

$$v(x,z) = 10^{-3}[1 - 10^2(x^2 + z^2)]$$

Find the volume of ice V moved through the xz-plane in a year (distances are in kilometers).

21. Let D_{xyz} be the parallelopiped bounded by the planes $x - y + z = 2$, $x - y + z = 3$, $x + 2y = -1$, $x + 2y = 1$, $x - z = 0$, $x - z = 2$. Evaluate

$$\iiint\limits_{D_{xyz}} e^{x-y+z}\, dV$$

22. Consider the region D in the first octant enclosed by the six surfaces

$$ay = x^3,\ by = x^3,\ cx = y^3,\ dx = y^3,\ z = 0,\ z = 1$$

where a, b, c, d are constants with $0 < a < b$, $0 < c < d$. Show that the volume of D equals $\frac{1}{2}\left(\sqrt{b} - \sqrt{a}\right)\left(\sqrt{d} - \sqrt{c}\right)$.

23. (a) The density of a spherical star of radius b depends on the distance $r = \sqrt{x^2 + y^2 + z^2}$ from the centre according to $\rho = f(r)$, where f is a positive, continuous function of one variable. Write the mass M of the star as a triple integral. Then show that

$$M = 4\pi \int_0^b r^2 f(r)\, dr$$

(b) The density of a spherical star of radius b is proportional to $\frac{b^3}{b^3 + r^3}$, where r is the distance to the centre. At what points does the density equal the average density?

24. A spherical star of radius b has a core of radius $\frac{1}{2}b$ with constant density ρ_0 kg/m^3. The density of the outer shell is proportional to $\frac{1}{r}$, where r is the distance to the centre. If the density is a continuous function of r, for $0 \le r \le b$, find the total mass of the star.

25.* The tetrahedron with vertices $(0,0,0)$, $(a,0,0)$, $(0,b,0)$ and $(0,0,c)$ is to be sliced into n pieces of equal volume by planes parallel to the inclined face. Where should the slices be made?

26.* Calculate the average distance of the point (0,0,c), where $c \ge 1$, from the set of all points in the solid sphere $x^2 + y^2 + z^2 \le 1$.

27.* A 3-sphere of radius b in $\mathbb{R}^4$ is defined by the equation

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = b^2$$

Find the "volume" enclosed by a 3-sphere of radius b.

28.* Find the volume of the region in $\mathbb{R}^3$ which is inside of **all** of the cylinders $x^2 + y^2 = 1$, $x^2 + z^2 = 1$, and $y^2 + z^2 = 1$. *Hint:* The region is *not* a sphere!

Appendix A

Implicitly Defined Functions

Implicit Differentiation

An equation of the form
$$f(x, y) = 0 \tag{A.1}$$
defines a relationship between the two variables x and y. If
$$y = g(x)$$
is a solution of equation (A.1), i.e.
$$f(x, g(x)) = 0 \tag{A.2}$$
for all x in some interval I, we say that the function g is **defined implicitly** by equation (A.1).

e.g. the functions $y = \sqrt{1 - x^2}$ and $y = -\sqrt{1 - x^2}$ are defined implicitly by the equation $x^2 + y^2 - 1 = 0$.

In general, given an equation of the form (A.1), it is not possible to solve for y in terms of x to obtain the function $g(x)$ explicitly. However it is easy to calculate the derivatives of g by differentiating equation (A.2) with respect to x, a process referred to as **implicit differentiation**. In this way one can find the linear approximation and second degree Taylor polynomial of g at a suitable reference point.

EXAMPLE 1 The equation
$$y^3 - y + x = 0 \tag{3}$$
defines y implicitly as a function of x, $y = g(x)$, with $g(0) = 1$. Find the linear approximation and second degree Taylor polynomial of g at the point $x = 0$.

Solution: Differentiate equation (3) with respect to x, treating y as a function of x:
$$3y^2 \frac{dy}{dx} - \frac{dy}{dx} + 1 = 0. \tag{4}$$

Evaluate this at the point $(x, y) = (0, 1)$, obtaining $\dfrac{dy}{dx} = -\frac{1}{2}$, and hence $g'(0) = -\frac{1}{2}$. Since $g(0) = 1$, the linear approximation of g at 0 is

$$L_0(x) = 1 - \frac{1}{2}x.$$

Differentiate equation (4) with respect to x:

$$3y^2\frac{d^2y}{dx^2} + 6y\left(\frac{dy}{dx}\right)^2 - \frac{d^2y}{dx^2} = 0.$$

Evaluate this at the point $(x, y) = (0, 1)$ obtaining $\dfrac{d^2y}{dx^2} = -\frac{3}{4}$, and hence $g''(0) = -\frac{3}{4}$. The second degree Taylor polynomial of g at 0 is

$$P_{2,0}(x) = 1 - \frac{1}{2}x - \frac{3}{8}x^2.$$

In this way, we can obtain information about the implicitly defined function g:

$$g(x) \approx 1 - \frac{1}{2}x - \frac{3}{8}x^2,$$

for x sufficiently close to 0.

EXERCISE 1

The equation

$$xy - \sin y = 0$$

defines y implicitly as a function of x, $y = g(x)$, with $g(0) = \pi$. Find the linear approximation of g at $x = 0$.

If the function $y = g(x)$ is defined implicitly by the equation $f(x, y) = 0$, where f has continuous partials, one can derive a formula for $g'(x)$ in terms of the partial derivatives of f. We have

$$f(x, g(x)) = 0$$

for all x in some interval. This equation states that the composite function $f(x, g(x))$ is the zero function. Thus

$$\frac{d}{dx}f(x, g(x)) = 0$$

Use the Chain Rule to expand this derivative, obtaining

$$f_x(x, g(x)) + f_y(x, g(x))g'(x) = 0$$

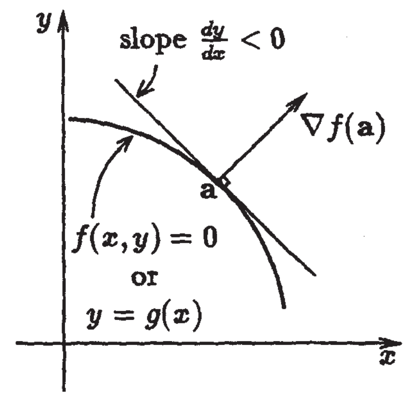
If $f_y(x, g(x)) \neq 0$, we can solve for $g'(x)$,

$$g'(x) = -\frac{f_x(x, g(x))}{f_y(x, g(x))}$$

It is not necessary to memorize this formula. What is of interest is its structure. In Leibniz notation, this equation reads

$$\frac{dy}{dx} = -\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}}$$

The minus sign if puzzling — one cannot think of "canceling the ∂f's", as is sometimes possible in single variable calculus.
There is a simple geometrical explanation, however. In the diagram, $\dfrac{\partial f}{\partial x} > 0$ and $\dfrac{\partial f}{\partial y} > 0$, based on the direction of the gradient vector, but $\dfrac{dy}{dx} < 0$, based on the slope of the tangent line.



EXERCISE 2 The equation

$$f(x, y) = 0$$

defines y implicitly as a function of x, $y = g(x)$. If $f(-1, 3) = 0$ and $\nabla f(-1, 3) = (3, 5)$, find $g'(-1)$. Assume that f has continuous partial derivatives.

Generalization

A function $g(x, y)$ can be defined implicitly by an equation of the form $f(x, y, z) = 0$ One can use implicit differentiation to calculate the partial derivatives of g. We assume that f has continuous partial derivatives.

EXAMPLE 2 The equation $f(x, y, z) = 0$ determines z implicitly as a function of x and y, $z = g(x, y)$. If $f(2, -1, 1) = 0$, and $\nabla f(2, -1, 1) = (4, -6, 2)$, find the linear approximation of g at $(-1, 1)$.

Solution: We have

$$f(x, y, g(x, y)) = 0 \tag{5}$$

for all (x, y) in some subset of $\mathbb{R}^2$. Differentiate equation (5) with respect to x, treating y as a constant:

$$\frac{\partial}{\partial x} f(x, y, g(x, y)) = 0$$

Expand the left side using the Chain Rule

$$f_x(x, y, g(x, y))(1) + f_y(x, y, g(x, y))(0) + f_z(x, y, g(x, y))g_x(x, y) = 0$$

Evaluate at $(x, y) = (2, -1)$, with $g(2, -1) = 1$:

$$f_x(2, -1, 1) + f_z(2, -1, 1)g_x(2, -1) = 0$$

Since $\nabla f(2, -1, 1) = (4, -6, 2)$, we obtain

$$4 + 2g_x(2, -1) = 0$$

and so

$$g_x(2, -1) = -2$$

Similarly one can show that

$$g_y(2, -1) = 3$$

The linear approximation of g at $(2, -1)$ is:

$$L_{(2,-1)}(x, y) = 1 - 2(x - 2) + 3(y + 1)$$

EXERCISE 3 Referring to Example 2, show that $g_y(2, -1) = 3$.

The Implicit Function Theorem

Above we showed that the derivatives of a function $y = g(x)$ that is defined implicitly by an equation $f(x, y) = 0$, can be calculated in a routine manner, even though the function g cannot be solved for explicitly.

In this section we show how to obtain more information about the set of points (x, y) which satisfy an equation $f(x, y) = 0$, called the **null set** of f, and denoted by $N(f)$:

$$N(f) = \{(x, y) \mid f(x, y) = 0\}$$

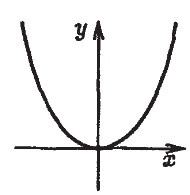
This set is simply the level curve of f which corresponds to the constant value[1] 0.

We begin by considering a number of simple examples which illustrate that it is difficult to make any general statements about the null set of f, even if f is a "well-behaved" function. In all the examples, f is a polynomial function, and hence has continuous partial derivatives of all orders.

(i) $f(x, y) = x^2 - y$

$N(f)$ is the graph of a differentiable function

$$y = g(x) = x^2$$

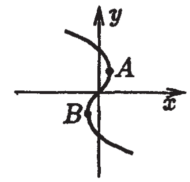


No exceptional points

[1]It is not important that we have used 0 as the constant value, since the level set $f(x, y) = k$, is the null set of the function g defined by $g(x, y) = f(x, y) - k$.

(ii) $f(x, y) = y^3 - y - x$

$N(f)$ is a smooth curve, which is not the graph of a function $y = g(x)$.



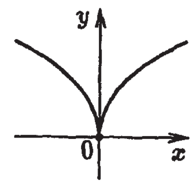
A and B are exceptional points

(iii) $f(x, y) = x^2 - y^3$

$N(f)$ is the graph of a non-differentiable function

$$y = g(x) = x^{2/3}$$

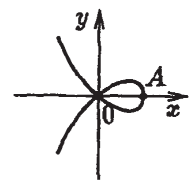
Note that $g'(0)$ does not exist.



0 is an exceptional point

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

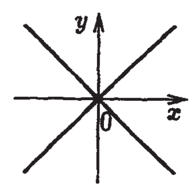
$N(f)$ is a self-intersecting curve (it could be the path of an electron in a magnetic field).



0 and A are exceptional points

(v) $f(x, y) = x^2 - y^2$

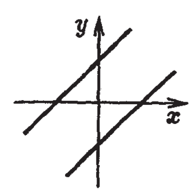
$N(f)$ consists of two intersecting curves.



0 is an exceptional point

(vi) $f(x, y) = (x - y)^2 - 1$

$N(f)$ consists of two disjoint curves.



No exceptional points

(vii) $f(x, y) = x^2 + y^2$

$N(f)$ is a single point $(0, 0)$.

(viii) $f(x, y) = x^2 + y^2 + 1$.

$N(f)$ is the empty set, i.e., 0 does not belong to the range of f.

REMARK

In general, for a given x-value, the equation $f(x, y) = 0$ does not have a unique solution for y, and may not have any solution. However, by studying the sketches, we see that apart from a few exceptional points, for each point $(a, b) \in N(f)$ there is a neighborhood of (a, b) such that when restricted to this neighborhood, $N(f)$ is the graph of a differentiable function $y = g(x)$. The function $y = g(x)$ represents the unique solution of the equation $f(x, y) = 0$ in this neighborhood.

The question is: how can we locate the exceptional points if we don't have a picture of the null set $N(f)$?

The answer is: by studying the gradient vector ∇f. Here's how:

If a level set $f(x, y) = 0$ is a smooth curve, and (a, b) lies on the curve (i.e. $f(a, b) = 0$), then $\nabla f(a, b)$ is normal to the tangent line to the curve at (a, b). Thus, at the exceptional points A and B in example (ii), and A in example (iv), at which the tangent line is vertical, $\nabla f = (f_x, 0)$ i.e. $f_y = 0$. At the exceptional point $(0, 0)$ in examples (iii), (iv), (v) and (vii), where the level set $f(x, y) = 0$ is not a smooth curve, we have $\nabla f = (0, 0)$, as can be verified explicitly (exercise).

The examples thus suggest that if $f_y(a, b) \neq 0$, then the level set $f(x, y) = 0$ is the graph of a function $y = g(x)$ in some neighborhood of (a, b), or equivalently that the equation $f(x, y) = 0$ has a unique solution $y = g(x)$. We now state this very important theorem.

THEOREM 1 **(Implicit Function Theorem)**

Let $f(x, y) \in C^1$ in a neighborhood of (a, b). If $f(a, b) = 0$ and $f_y(a, b) \neq 0$, then there exists a neighborhood of (a, b) in which the equation $f(x, y) = 0$ has a unique solution for y in terms of x, $y = g(x)$, where g has a continuous derivative.

REMARK

The roles of the variables x and y can be interchanged. If the hypothesis $f_y(a, b) \neq 0$ is replaced by $f_x(a, b) \neq 0$, the conclusion is that the equation $f(x, y) = 0$ has a unique solution for x,

$$x = h(y)$$

The theorem and comment lead to the following:

COROLLARY 2

If $f(x,y) \in C^1$, and

$$f(a,b) = 0, \qquad \nabla f(a,b) \neq (0,0)$$

(i.e. at least one partial derivative non-zero at (a,b)), then near the point (a,b), the equation $f(x,y) = 0$ describes a smooth curve, whose tangent line at (a,b) is orthogonal to $\nabla f(a,b)$. If $f_y(a,b) \neq 0$ then the curve can be written uniquely in the form

$$y = g(x)$$

and if $f_x(a,b) \neq 0$, it can be written uniquely in the form

$$x = h(y)$$

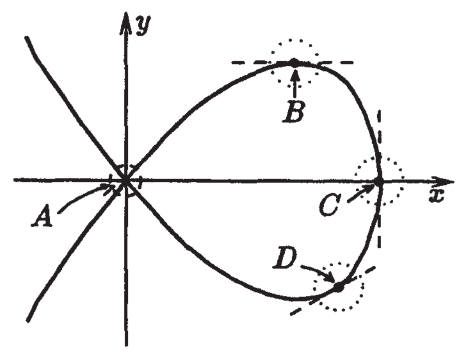
In the sketch below, we illustrate the Implicit Function Theorem using the function $f(x,y) = -x^2 + x^3 + y^2$ in example (iv).

A: $\nabla f(0,0) = (0,0)$.

Not a smooth curve in this neighborhood. There is not a unique solution.

B: $\nabla f(\frac{2}{3}, \frac{2}{3\sqrt{3}} = (0, \frac{4}{3\sqrt{3}}$.

Smooth curve in this neighborhood. Unique solution $y = g(x)$.



C: $\nabla f(1,0) = (1,0)$.

Smooth curve in this neighborhood. Unique solution $x = h(y)$.

D: $\nabla f(\frac{3}{4}, -\frac{3}{8}) = (\frac{3}{16}, -\frac{3}{4})$.

Smooth curve in this neighborhood. Unique solution $y = g(x)$ and $x = h(y)$.

EXERCISE 4

a. Prove that the equation $f(x,y) = 2x^2 - 2y^2 + y^4 = 0$ has a unique solution $y = g(x)$ near the point $\left(\frac{\sqrt{7}}{4\sqrt{2}}, \frac{1}{2}\right)$.

b. At what points is the tangent line to the curve $f(x,y) = 0$ horizontal/vertical?

c. Use b. to sketch the set defined by $f(x,y) = 0$.

It is a self-intersecting curve with a familiar shape.

Generalization

The considerations of this section can be applied to an equation of the form

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

The geometric interpretation is in terms of surfaces in $\mathbb{R}^3$.

We first state the Implicit Function Theorem and its corollary for functions of three-variables $f(x, y, z)$.

THEOREM 3 **(Implicit Function Theorem)**

Let $f(x, y, z) \in C^1$ in a neighborhood of (a, b, c). If $f(a, b, c) = 0$ and $f_z(a, b, c) \neq 0$, then there exists a neighborhood of (a, b, c) in which the equation $f(x, y, z) = 0$ has a unique solution for z in terms of x and y, $z = g(x, y)$, where $g \in C^1$.

COROLLARY 4 If $f(x, y, z)$ has continuous partial derivatives, and

$$f(a, b, c) = 0, \qquad \nabla f(a, b, c) \neq (0, 0, 0)$$

(i.e. at least one partial derivative is non-zero at (a, b, c)), then near the point (a, b, c), the equation $f(x, y, z) = 0$ describes a smooth surface in $\mathbb{R}^3$ whose tangent plane at (a, b, c) is orthogonal to $\nabla f(a, b, c)$.

If $f_z(a, b, c) \neq 0$, then the surface can be described uniquely in the form

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

near the point (a, b, c). In general, however, the equation $f(x, y, z) = 0$ will not be the graph $z = g(x, y)$ of one function g.

EXAMPLE 3 $f(x, y, z) = x^2 + y^2 + z^2 - 1 = 0$ represents a sphere, and can thus be described by the graphs of two functions,

$$z = \sqrt{1 - x^2 - y^2} \quad \text{and} \quad z = -\sqrt{1 - x^2 - y^2}$$

REMARK

When applying the Implicit Function Theorem it is easy to remember which partial derivative of f must be non-zero: it is the partial derivative with respect to the variable for which one wishes to solve.

EXAMPLE 4 Prove that the equation

$$F(x, y, z) = ye^z + xz - x^2 - y^2 = 0$$

has a unique solution for x in terms of y and z in a neighborhood of $(0, 2, \ln 2)$.

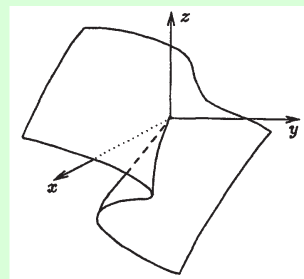
Solution: F has continuous partials for all $(x, y, z) \in \mathbb{R}^3$ by inspection. In addition $F(0, 2, \ln 2) = 0$. The essential condition is that $\dfrac{\partial F}{\partial x}(0, 2, \ln 2) \neq 0$.

This is easily verified, since $\dfrac{\partial F}{\partial x} = z - 2x$.

EXAMPLE 5 The equation

$$f(x, y, z) = z^3 - xz + y = 0 \qquad \text{(A.3)}$$

describes a smooth surface with a "fold", like a wave on the point of breaking. Show that the curve $\mathbf{x} = (3t^2, 2t^3, t)$, $t \in \mathbb{R}$ lies on the surface and that the tangent plane is vertical at each point of this curve.



Solution: Firstly, show that $f(x, y, z)$ is zero along the given curve:

$$f(3t^2, 2t^3, t) = t^3 - (3t^2)(t) + 2t^3 = 0$$

for all $t \in \mathbb{R}$. Thus, the curve lies in the surface.

The gradient vector is $\nabla f(x, y, z) = (-z, 1, 3z^2 - x)$.

Evaluate this vector on the curve:

$$\nabla f(3t^2, 2t^3, t) = (-t, 1, 3t^2 - 3t^2) = (-t, 1, 0)$$

This shows that ∇f is parallel to the xy-plane at points on the curve. Since ∇f is orthogonal to the tangent plane of the surface $f(x, y, z) = 0$, it follows that the tangent plane of the surface $f(x, y, z) = 0$ is vertical at points on the given curve.

REMARK

The fact that equation (A.3) does describe a smooth surface follows from the fact that (A.3) can be solved for y by inspection: $y = g(x, z) = xz - z^3$ and that g has continuous partials.

EXERCISE 5 In order to verify the shape of the surface given by equation (A.3), sketch some typical cross-sections $x = a$, given by $z^3 - az + y = 0$ in the cases $a > 0$, $a = 0$, $a < 0$.

Appendix B

Coordinate Systems

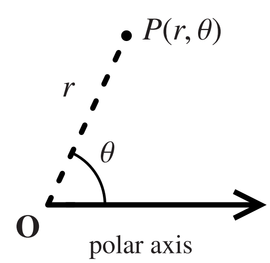
A **coordinate system** is a system for representing the location of a point in a space by an ordered n-tuple. We call the elements of the n-tuple the **coordinates** of the point.

We are used to using the Cartesian coordinate system in which the location of the point is represented by the directed distance from a set of perpendicular axes which all intersect at a point O. However, you may also be used to other coordinate systems. For example, the geographic coordinate system represents location on the earth by longitude, latitude and altitude.

We will now look at three other important coordinate systems: polar coordinates, cylindrical coordinates, and spherical coordinates.

B.1 Polar Coordinates

As in all coordinate systems, we must have a frame of reference for our coordinate system. So, in a plane we choose a point O called the **pole** (or origin). From O we draw a ray called the **polar axis**. Generally, the polar axis is drawn horizontally to the right to match the positive x-axis in Cartesian coordinates.



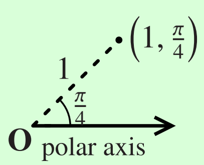
Let P be any point in the plane. We will represent the position of P by the ordered pair (r, θ) where $r \geq 0$ is the length of the line OP and θ is the angle between the polar axis and OP. We call r and θ the polar coordinates of P.

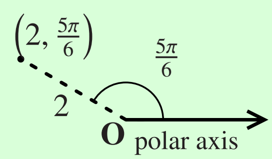
REMARKS

1. We assume, as usual, that an angle θ is considered positive if measured in the counterclockwise direction from the polar axis and negative if measured in the clockwise direction.
2. We represent the point O by the polar coordinates $(0, \theta)$ for any value of θ.
3. We are restricting r to be non-negative to coincide with the interpretation of r as distance. Many textbooks do not put this restriction on r.
4. Since we use the distance r from the pole in our representation, polar coordinates are suited for solving problems in which there is symmetry about the pole.

EXAMPLE 1 Plot the points $\left(1, \frac{\pi}{4}\right)$ and $\left(2, \frac{5\pi}{6}\right)$ in polar coordinates.

Solution:





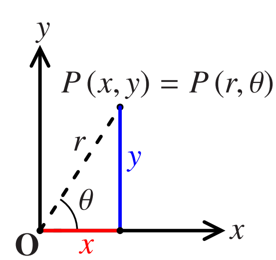
There is one important difference between polar coordinates and Cartesian coordinates. In Cartesian coordinates each point has a unique representation (x, y). However, observe that a point (r, θ) in polar coordinates can have infinitely many representations. In particular,

$$(r, \theta) = (r, \theta + 2\pi k), \quad k \in \mathbb{Z}$$

Relationship to Cartesian Coordinates

If we now place the pole O at the origin of the Cartesian plane and lie the polar axis along the positive x-axis, we can find a relationship between the coordinates of a point P in the two coordinate systems. In particular, we see from the diagram that

$$\begin{aligned} x &= r\cos\theta, & r &= \sqrt{x^2 + y^2} \\ y &= r\sin\theta, & \tan\theta &= \frac{y}{x} \end{aligned} \tag{B.1}$$



EXAMPLE 2 Convert the points $\left(2, -\frac{\pi}{3}\right)$ and $\left(1, \frac{3\pi}{4}\right)$ from polar coordinates to Cartesian coordinates.

Solution: We have $x = 2\cos -\frac{\pi}{3} = 1$ and $y = 2\sin -\frac{\pi}{3} = -\sqrt{3}$. Hence, the point is $(1, -\sqrt{3})$ in Cartesian coordinates.

We have $x = \cos \frac{3\pi}{4} = -\frac{1}{\sqrt{2}}$ and $y = \sin \frac{3\pi}{4} = \frac{1}{\sqrt{2}}$. So, the point has Cartesian coordinates $\left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)$.

EXAMPLE 3 Convert the point $(1, 1)$ from Cartesian coordinates to polar coordinates.

Solution: We have $x = 1$ and $y = 1$, so $r = \sqrt{1^2 + 1^2} = \sqrt{2}$ and $\tan\theta = 1$. Since x and y are both positive the point is in quadrant 1, and hence

$$\theta = \frac{\pi}{4} + 2\pi k, \quad k \in \mathbb{Z}$$

Therefore, we get the polar coordinate representations $\left(\sqrt{2}, \frac{\pi}{4} + 2\pi k\right)$, $k \in \mathbb{Z}$.

Often we do not need to find all possible polar representations for a point. Thus, we further restrict ourselves to a range of θ (such as $0 \leq \theta < 2\pi$ or $-\pi < \theta \leq \pi$) which gives unique representation.

EXAMPLE 4 Convert the point $(-1, \sqrt{3})$ from Cartesian coordinates to polar coordinates with $0 \leq \theta < 2\pi$.

Solution: We have $x = -1$ and $y = \sqrt{3}$, so $r = \sqrt{(-1)^2 + (\sqrt{3})^2} = 2$ and $\tan\theta = -\sqrt{3}$. Since θ is in the second quadrant we get $\theta = \frac{2}{3}\pi$. Hence, the point has polar representation $\left(2, \frac{2}{3}\pi\right)$.

REMARK

The equation $\tan\theta = \frac{y}{x}$ does not uniquely determine θ, since over $0 \leq \theta \leq 2\pi$ each value of $\tan\theta$ occurs twice. One must be careful to choose the θ which lies in the correct quadrant.

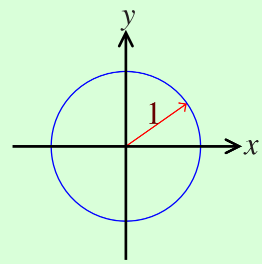
Graphs in Polar Coordinates

The graph of an explicitly defined polar equation $r = f(\theta)$ or $\theta = f(r)$, or an implicitly defined polar equation $f(r, \theta) = 0$, is a curve that consists of all points that have at least one polar representation (r, θ) that satisfies the equation of the curve.

EXAMPLE 5

Sketch the polar equation $r = 1$.

Solution: This is the curve which consists of all points $(r, \theta) = (1, \theta)$, $\theta \in \mathbb{R}$. Observe that this is all points that have distance 1 from the origin. Hence, we get a circle of radius 1.



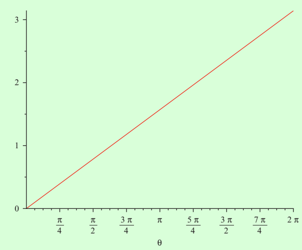
EXERCISE 1

Sketch the polar equation $\theta = \dfrac{\pi}{4}$.

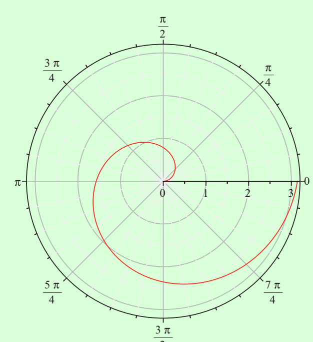
EXAMPLE 6

Sketch the polar equation $r = \dfrac{1}{2}\theta$, $0 \leq \theta \leq 2\pi$.

Solution: One way to try to sketch a curve is to make a table of values and plot the points for various θ, however this is quite tedious. Instead, let's consider sketching the curve as if it was given in Cartesian coordinates in the $r\theta$-plane:



Essentially we have created a table of infinitely many values which allows us to see how r grows as θ increases from 0 to 2π. Finally, we sketch the given curve in the xy-plane where r is the distance to the origin and θ is the angle measured counterclockwise from the x-axis. We see that the distance from the origin grows linearly as we increase the angle; we get a spiral:

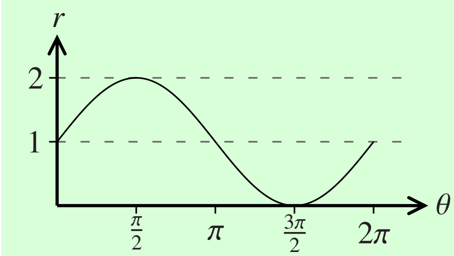


REMARK

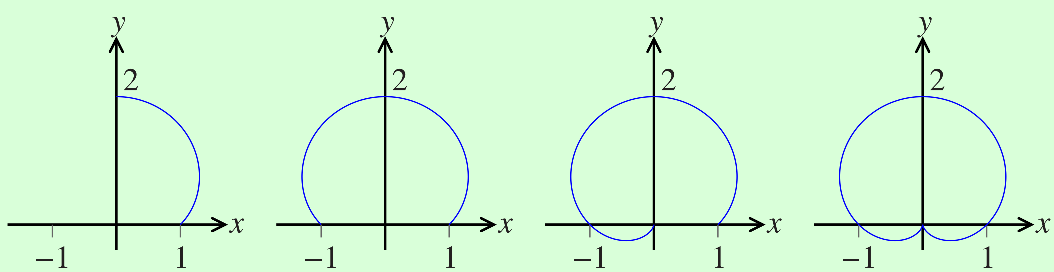
The polar equation $r = e^{\theta}$ gives a **logarithmic spiral** which often appears in nature.

EXAMPLE 7 Sketch the polar equation $r = 1 + \sin\theta$.

Solution: To sketch this equation we first sketch the curve in Cartesian coordinates in the $r\theta$-plane and use this graph to plot points in the xy-plane.

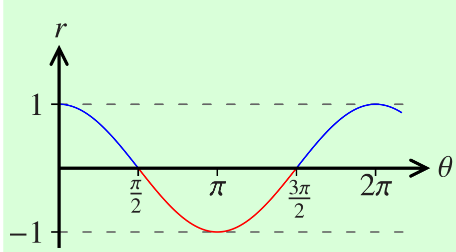


Observe from the diagram that as θ increases from 0 to $\frac{\pi}{2}$ the radius increases from 1 to 2. Then when θ increases from $\frac{\pi}{2}$ to π the radius decreases from 2 to 1. As θ increases from π to $\frac{3\pi}{2}$ we get the radius decreases from 1 to 0, and as θ increases from $\frac{3\pi}{2}$ to 2π the radius increases from 0 to 1. Each of these steps are shown below. The final curve is called a **cardioid**.

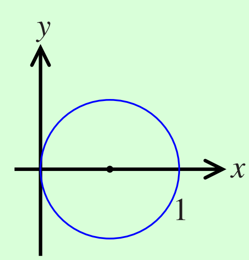


EXAMPLE 8 Sketch the polar equation $r = \cos\theta$.

Solution: We first sketch the curve in Cartesian coordinates in the $r\theta$-plane.



We see that as θ increases from 0 to $\frac{\pi}{2}$ the radius decreases from 1 to 0. For values of θ from $\frac{\pi}{2}$ to $\frac{3\pi}{2}$ the radius is negative, thus we do not draw any points since we have made the restriction that $r \geq 0$. As θ moves from $\frac{3\pi}{2}$ to 2π the radius increases from 0 to 1.



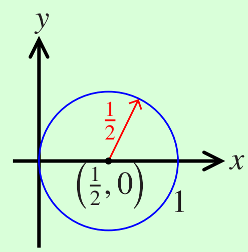
EXERCISE 2 Sketch the polar equations $r = \sin\theta$ and $r = 1 - 2\cos\theta$.

We have seen above that we can use equations (B.1) to convert points between the coordinate systems. Thus, we can also use these equations to convert equations of curves between the two coordinate systems.

EXAMPLE 9 Convert the equation $r = \cos\theta$ to Cartesian coordinates.

Solution: Since $r^2 = x^2 + y^2$ and $x = r\cos\theta$, we get

$$\begin{aligned} r &= \cos\theta \\ r^2 &= r\cos\theta \\ x^2 + y^2 &= x \\ \left(x - \frac{1}{2}\right)^2 + y^2 &= \frac{1}{4}. \end{aligned}$$



Which is a circle of radius $\frac{1}{2}$ centered at $\left(\frac{1}{2}, 0\right)$ as we drew in Example 8.

EXAMPLE 10 Convert the equation of the curve $(x^2 + y^2)^{3/2} = 2xy$ to polar coordinates.

Solution: Since $x = r\cos\theta$ and $y = r\sin\theta$ we get

$$\begin{aligned} (x^2 + y^2)^{3/2} &= 2xy \\ r^3 &= 2(r\cos\theta)(r\sin\theta) \\ r^3 &= r^2 \sin 2\theta \\ r &= \sin 2\theta \end{aligned}$$

Notice that the last simplification is only valid since the pole $r = 0$, is still included in the graph (the case where $\theta = \pi$).

Observe that since we have the restriction $r \geq 0$ we must also have $\sin 2\theta \geq 0$. Hence, we find that a domain of the function is

$$0 \leq \theta \leq \frac{\pi}{2}, \quad \pi \leq \theta \leq \frac{3\pi}{2}$$

EXERCISE 3 Convert the equation of the curve $x^2 - y^2 = 1$ to polar coordinates.

Area in Polar Coordinates

We now wish to derive the formula for computing area between curves in Polar coordinates. Clearly this will be a little different than before as it does not make sense to use rectangles to find our area. In Polar coordinates, it is natural to use sectors of a circle.

Recall that if θ_1 and θ_2, $\theta_2 > \theta_1$, are two angles in a circle of radius r, then the area between them is

$$\frac{\theta_2 - \theta_1}{2\pi} \cdot \pi r^2 = \frac{1}{2} r^2 (\theta_2 - \theta_1)$$

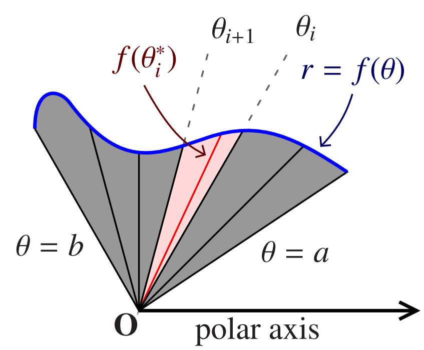
We now derive the area as before. We divide the region bounded by $\theta = a$, $\theta = b$ and $r = f(\theta)$ into subregions $\theta_0, \ldots, \theta_n$ of equal difference $\Delta\theta$, then for each subregion θ_i, $0 \leq i < n$ we pick some point θ_i^* with $\theta_i \leq \theta_i^* \leq \theta_{i+1}$. We then form the sector between θ_i and θ_{i+1} with radius $f(\theta_i^*)$. The area of this sector is

$$\frac{1}{2}[f(\theta_i^*)]^2 \Delta\theta$$

Hence, the area is approximately

$$\sum_{i=0}^{n-1} \frac{1}{2}[f(\theta_i^*)]^2 \Delta\theta$$

Thus, as we let the number of subdivisions go to infinity and hence letting each of the $\Delta\theta_i$ tend to 0 we get



$$A = \lim_{\|\Delta\theta_i\| \to 0} \sum_{i=0}^{n-1} \frac{1}{2}[f(\theta_i^*)]^2 \Delta\theta = \int_a^b \frac{1}{2}[f(\theta)]^2 \, d\theta$$

EXAMPLE 11 Find the area inside the circle $r = a$.

Solution: We need θ to range from 0 to 2π to make the whole circle so we have

$$A = \int_0^{2\pi} \frac{1}{2} a^2 \, d\theta = \frac{1}{2} a^2 [2\pi - 0] = \pi a^2$$

EXERCISE 4 Find the area inside the lemniscate $r = 2\sqrt{\sin 2\theta}$.

ALGORITHM

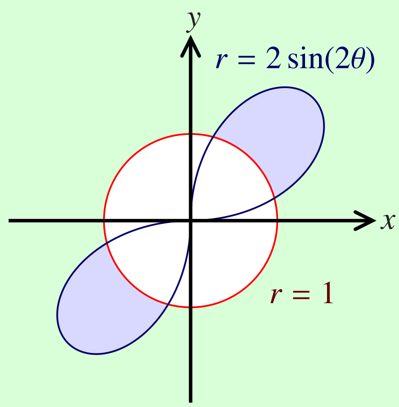
To find the area between two curves in Polar coordinates, we use the same method we used for doing this in Cartesian coordinates.

1. Find the points of intersections.
2. Graph the curves and split the desired region into easily integrable regions.
3. Integrate.

EXAMPLE 12 Find the area inside $r = 2\sin(2\theta)$, but outside $r = 1$.

Solution:

Setting the curves equal to each other we get $1 = 2\sin(2\theta)$, hence $2\theta = \frac{\pi}{6}$ or $2\theta = \frac{5\pi}{6}$. Therefore, we want to integrate over the region $\frac{\pi}{12}$ to $\frac{5\pi}{12}$. To find the shaded area, we will find the area inside the lemniscate in the first quadrant and subtract off the area of the region that is inside both the circle and the lemniscate. Finally, we will multiply by 2 for the symmetric region in the third quadrant. We get



$$\begin{aligned} A &= 2\left(\int_{\pi/12}^{5\pi/12} \frac{1}{2}(2\sin(2\theta))^2\, d\theta - \int_{\pi/12}^{5\pi/12} \frac{1}{2}(1)^2\, d\theta\right) \\ &= \cdots = \frac{\pi}{3} + \frac{\sqrt{3}}{2} \end{aligned}$$

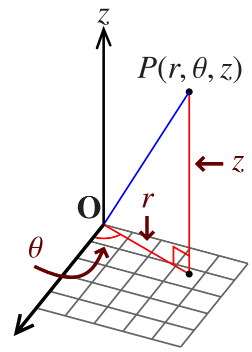
REMARK

Finding points of intersection can be tricky, especially at the pole/origin which does not have a unique representation: $(0, \theta)$ for any θ represents the origin, so simply setting expressions equal to each other may 'miss' that point. It is essential to sketch the region when finding points of intersection.

EXERCISE 5 Find the area between the curves $r = \cos\theta$ and $r = \sin\theta$.

B.2 Cylindrical Coordinates

Observe that we can extend polar coordinates to 3-dimensional space by introducing another axis, called the **axis of symmetry**, through the pole perpendicular to the polar plane. We then represent any point P in the space by the cylindrical coordinates (r, θ, z) where r and θ are as in polar coordinates and z is the height above (or below) the polar plane. Thus, as in Polar coordinates, we have the restrictions $r \geq 0$, $0 \leq \theta < 2\pi$ (or $-\pi < \theta \leq \pi$).

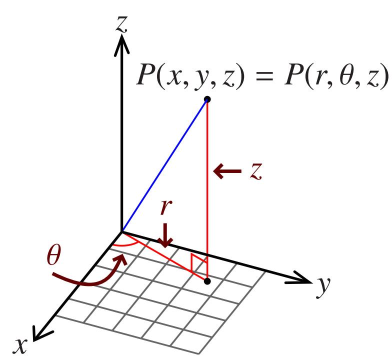


REMARK

Notation for cylindrical coordinates may vary from author to author. In particular, in the sciences they generally use the Standard ISO 31-11 notation which gives the cylindrical coordinates as (ρ, ϕ, z).

If we place the pole at the origin and the polar axis along the positive x-axis as in polar coordinates and place the axis of symmetry along the z-axis we then can relate a point P in cylindrical and Cartesian coordinates by

$$\begin{aligned} x &= r\cos\theta, & r &= \sqrt{x^2+y^2} \\ y &= r\sin\theta, & \tan\theta &= \frac{y}{x} \\ z &= z, & z &= z \end{aligned} \tag{B.2}$$



REMARK

Cylindrical coordinates are useful when there is symmetry about an axis. Thus, it is sometimes desirable to lie the polar axis and axis of symmetry along different axes.

EXAMPLE 1

Convert $(2,0,0)$ and $(0,\pi,2)$ from cylindrical coordinates to Cartesian coordinates.

Solution: The first point has coordinates $r = 2$, $\theta = 0$ and $z = 0$. Hence, $x = 2\cos 0 = 2$, $y = 2\sin 0 = 0$ and $z = 0$, so we the point in Cartesian coordinates is $(2,0,0)$.

The second point has coordinates $r = 0$, $\theta = \pi$, and $z = 2$. Since, $r = 0$ we get $x = y = 0$ and so the point in Cartesian coordinates is $(0,0,2)$.

EXAMPLE 2

Convert $(1,1,3)$ and $(1,-\sqrt{3},1)$ from Cartesian coordinates to cylindrical coordinates.

Solution: We have $r = \sqrt{1^2+1^2} = \sqrt{2}$, $\tan\theta = 1$ which gives $\theta = \frac{\pi}{4}$ and $z = 3$. Thus, in cylindrical coordinates the point is $\left(\sqrt{2}, \frac{\pi}{4}, 3\right)$.

We have $z = 1$, $r = \sqrt{1^2 + (-\sqrt{3})^2} = 2$, $\tan\theta = \frac{-\sqrt{3}}{1}$ which gives $\theta = \frac{5\pi}{3}$ since θ is in the fourth quadrant. Hence, in cylindrical coordinates the point is $\left(2, \frac{5\pi}{3}, 1\right)$.

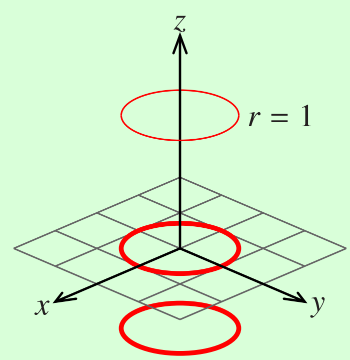
Graphs in Cylindrical Coordinates

As with functions $z = f(x,y)$, the graphs of functions $z = f(r,\theta)$, or more generally, $f(r,\theta,z) = 0$ are surfaces in $\mathbb{R}^3$.

EXAMPLE 3

Sketch the graph of $r = 1$ in cylindrical coordinates.

Solution: We know that $r = 1$ gives a circle of radius 1 in polar coordinates. Thus, in cylindrical coordinates we have a circle of radius 1 at any value of z. Hence, we have an infinite cylinder of radius 1.



EXERCISE 1

Sketch the graph of $z = r^2$ in cylindrical coordinates.

As we did in polar coordinates, we can also transform the equations of curves between the coordinates systems.

EXAMPLE 4

Convert the equation $z = r^2 \cos\theta$ to Cartesian coordinates.

Solution: Using (B.2) we get $z = x\sqrt{x^2 + y^2}$.

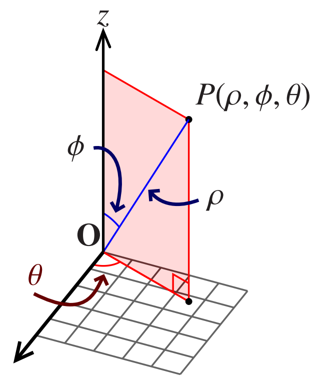
EXERCISE 2

Find the equation of $z = \dfrac{y}{\sqrt{x^2 + y^2}}$ in cylindrical coordinates.

B.3 Spherical Coordinates

In 2-dimensional space, we saw that polar coordinates were useful for problems which where symmetric about the origin. We now extend this idea to another 3-dimensional coordinate system called **spherical coordinates**.

As we did in cylindrical coordinates, we will use the pole O and polar axis from polar coordinates and draw another axis z perpendicular to the polar plane.
Let P be any point in 3-dimensional space. We will represent P by the coordinates (ρ, ϕ, θ) where $\rho \geq 0$ is the length of the line OP, θ is the same angle as in cylindrical coordinates, and ϕ is the angle between the positive z-axis and the line OP.



Since we are keeping the same interpretation of θ from cylindrical coordinates, it tells us the orientation of P around the z-axis. Therefore, we only want ϕ to indicate the "tilt" of the point with the z-axis. So, we restrict $0 \leq \phi \leq \pi$.

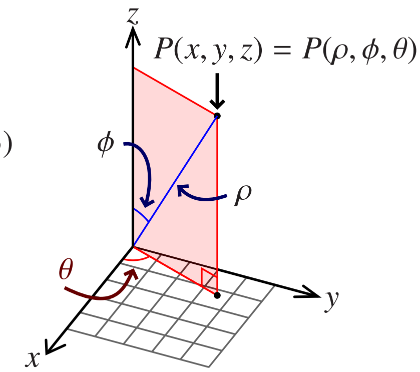
Thus, our restrictions in spherical coordinates are $\rho \geq 0$, $0 \leq \theta < 2\pi$ (or $-\pi < \theta \leq \pi$) and $0 \leq \phi \leq \pi$.

REMARK

The symbols used for spherical coordinates also vary from author to author as does the order in which they are written. In mathematics, it is not uncommon to find ρ replaced by r. The standard ISO 31-11 convention uses ϕ as the polar angle and θ as the angle with the positive z-axis. Therefore, it is very important to understand which notation is being used when reading an article.

From the diagram, we see that we can convert a point from Cartesian coordinates to spherical coordinates with the equations:

$$x = \rho \sin\phi \cos\theta, \qquad \rho = \sqrt{x^2+y^2+z^2}$$
$$y = \rho \sin\phi \sin\theta, \qquad \tan\theta = \frac{y}{x} \qquad \text{(B.3)}$$
$$z = \rho\cos\phi, \qquad \cos\phi = \frac{z}{\sqrt{x^2+y^2+z^2}}$$



EXAMPLE 1 Convert $\left(1, \frac{\pi}{4}, \frac{\pi}{4}\right)$ and $\left(1, \frac{\pi}{4}, \frac{5\pi}{4}\right)$ from spherical coordinates to Cartesian coordinates.

Solution: We get $x = \sin\frac{\pi}{4}\cos\frac{\pi}{4} = \frac{1}{2}$, $y = \sin\frac{\pi}{4}\sin\frac{\pi}{4} = \frac{1}{2}$, and $z = \cos\frac{\pi}{4} = \frac{1}{\sqrt{2}}$. Therefore, the point has Cartesian coordinates $\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{\sqrt{2}}\right)$.

We get $x = \sin\frac{\pi}{4}\cos\frac{5\pi}{4} = -\frac{1}{2}$, $y = \sin\frac{\pi}{4}\sin\frac{5\pi}{4} = -\frac{1}{2}$, and $z = \cos\frac{\pi}{4} = \frac{1}{\sqrt{2}}$. Therefore, the point has Cartesian coordinates $\left(-\frac{1}{2}, -\frac{1}{2}, \frac{1}{\sqrt{2}}\right)$.

EXAMPLE 2 Convert $\left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, \sqrt{3}\right)$ and $(-1, -1, -1)$ from Cartesian coordinates to spherical coordinates.

Solution: We have $\rho = \sqrt{\left(\frac{1}{\sqrt{2}}\right)^2 + \left(\frac{1}{\sqrt{2}}\right)^2 + \left(\sqrt{3}\right)^2} = 2$, $\tan\theta = 1 \Rightarrow \theta = \frac{\pi}{4}$ since θ is in the first quadrant and $\cos\phi = \frac{\sqrt{3}}{2} \Rightarrow \phi = \frac{\pi}{6}$. Hence, in spherical coordinates the point is $\left(2, \frac{\pi}{6}, \frac{\pi}{4}\right)$.

We get $\rho = \sqrt{(-1)^2 + (-1)^2 + (-1)^2} = \sqrt{3}$, $\tan\theta = 1 \Rightarrow \theta = \frac{5\pi}{4}$ since θ is the third quadrant, and $\cos\phi = \frac{-1}{\sqrt{3}}$. Thus, the point in spherical coordinates is $\left(\sqrt{3}, \arccos\frac{-1}{\sqrt{3}}, \frac{5\pi}{4}\right)$.

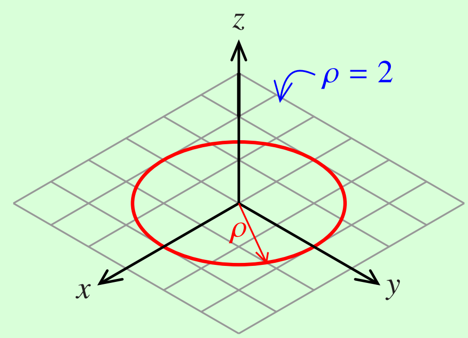
Observe from the above examples, how θ controls which quadrant the point is in (its rotation around the z-axis) and ϕ only controls whether the point will be above or below the xy-plane.

Graphs in Spherical Coordinates

As with cylindrical coordinates, the graph of a function $f(\rho, \phi, \theta) = 0$ in spherical coordinates gives a surface in $\mathbb{R}^3$.

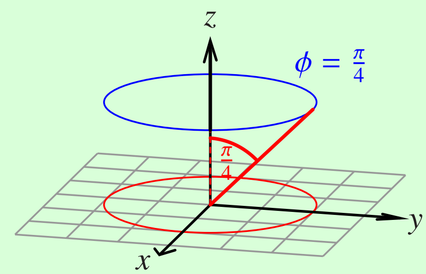
EXAMPLE 3 Sketch $\rho = 2$.

Solution: Observe that this is the graph with all points 2 units from the origin. Hence, it is a sphere of radius 2.



EXAMPLE 4 Sketch $\phi = \dfrac{\pi}{4}$.

Solution: First imagine a line which makes a $\frac{\pi}{4}$ angle with the positive z-axis. Since there is no restriction on θ, the graph of the surface will be this line rotated around the positive z-axis. Hence, we get a cone.



As with the other coordinate systems, we also want to convert equations between Cartesian and spherical coordinates.

EXAMPLE 5 Convert $\rho = \sin\phi\cos\theta$ to Cartesian coordinates.

Solution: We first multiply both sides of the equation by ρ to get

$$\rho^2 = \rho\sin\phi\cos\theta$$

Hence, we can apply (B.3) to get

$$\begin{aligned} x^2 + y^2 + z^2 &= x \\ (x - \frac{1}{2})^2 + y^2 + z^2 &= \frac{1}{4} \end{aligned}$$

EXAMPLE 6 Convert $z^2 = x^2 + y^2$ to spherical coordinates.

Solution: We have

$$\begin{aligned}\rho^2\cos^2\phi &= \rho^2\sin^2\phi\cos^2\theta + \rho^2\sin^2\phi\sin^2\theta\\ \cos^2\phi &= \sin^2\phi(\cos^2\theta + \sin^2\theta)\\ \tan^2\phi &= 1\end{aligned}$$

Thus, $\tan\phi = \pm 1$, so we get $\phi = \frac{\pi}{4}$ or $\phi = \frac{3\pi}{4}$. Observe that $\phi = \frac{\pi}{4}$ is the top-half of the cone (as in Example 4) and $\phi = \frac{3\pi}{4}$ is the bottom-half of the cone.

EXERCISE 1 Convert $x^2 + y^2 + z^2 = 2x$ to spherical coordinates.

Appendix B Problem Set

1. Convert the following points from Cartesian coordinates to polar coordinates with $0 \le \theta < 2\pi$.
   (a) $(-2, 2)$ (b) $(\sqrt{3}, -1)$
   (c) $(-1, -\sqrt{3})$ (d) $(2, 1)$
2. Convert the following points from polar coordinates to Cartesian coordinates.
   (a) $(2, \pi/3)$ (b) $(3, 5\pi/6)$
   (c) $(3, 2\pi/3)$ (d) $(2, -\pi/6)$
3. For each of the indicated regions in polar coordinates, sketch the region and find the area.
   (a) The region enclosed by $r = \sin\theta$
   (b) The region enclosed by $r = \cos 2\theta$
4. For each of the indicated regions in polar coordinates, sketch the region and find the area.
   (a) Inside both $r = 1 + 1\sin\theta$ and $r = 1 - 1\sin\theta$
   (b) Inside $r = \sin\theta$ and outside $r = \sin 2\theta$
5. For each of the indicated regions in polar coordinates, sketch the region and find the area.
   (a) The region enclosed by $r = \sin 3\theta$
   (b) Inside both $r = 2 + 2\cos\theta$ and $r = 2 - 2\cos\theta$
6. Convert the following equations in Cartesian coordinates to cylindrical coordinates.
   (a) $z = \sqrt{2x^2 + 2y^2}$
   (b) $x = y$
   (c) $z^2 = x^2 - y^2$
7. Convert the following equations in Cartesian coordinates to cylindrical coordinates.
   (a) $z = x^2 + y^2$
   (b) $1 = x^2 - y^2$
8. Convert the following equations in Cartesian coordinates to spherical coordinates.
   (a) $x^2 + y^2 = 4$
   (b) $x^2 + y^2 + z^2 = 2x$
   (c) $z = -\sqrt{x^2 + y^2}$
   (d) $z^2 = x^2 - y^2$
9. Convert the following equations in Cartesian coordinates to spherical coordinates.
   (a) $x = y$
   (b) $(x^2 + y^2 + z^2)^2 = z$
10. For each of the following regions in $\mathbb{R}^3$, given in Cartesian coordinates,
    (i) Give a description in spherical coordinates.
    (ii) Give a description in cylindrical coordinates.

    (a) $C = \{(x, y, z) : z \ge \sqrt{x^2 + y^2},\ x^2+y^2+(z-1)^2 \le 1\}$ (ice cream cone)
    (b) $L = \{(x, y, z) : z \ge \sqrt{x^2 + y^2},\ x^2 + y^2 + z^2 \le 2\}$ (licked ice cream cone)
    (c) $R = \{(x, y, z) : z \ge \sqrt{x^2 + y^2},\ z \le 1\}$ (really licked ice cream cone)

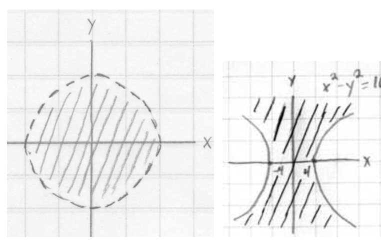
Appendix C

Answers to Mid-Section Exercises

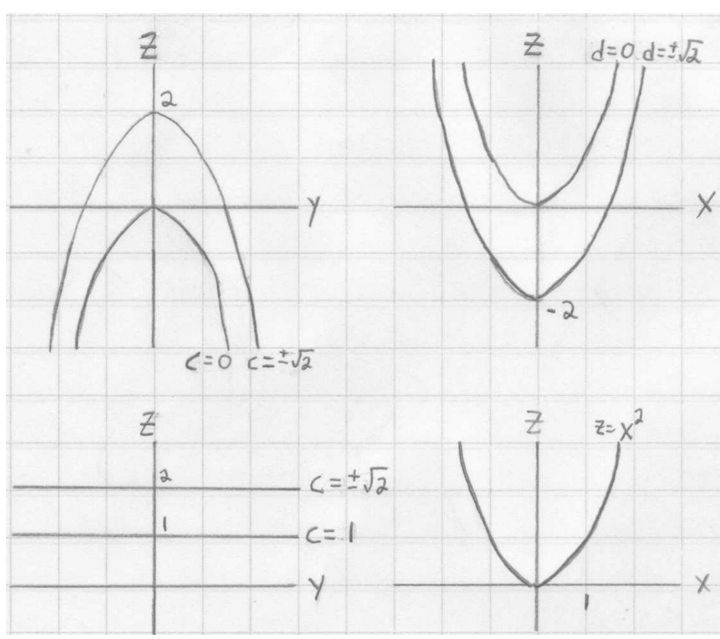
Answers to Chapter 1

1.1 **Exercise 1:** a) The domain of f is $1 - x^2 - y^2 > 0 \Rightarrow x^2 + y^2 < 1$. The range is $z \leq 0$.

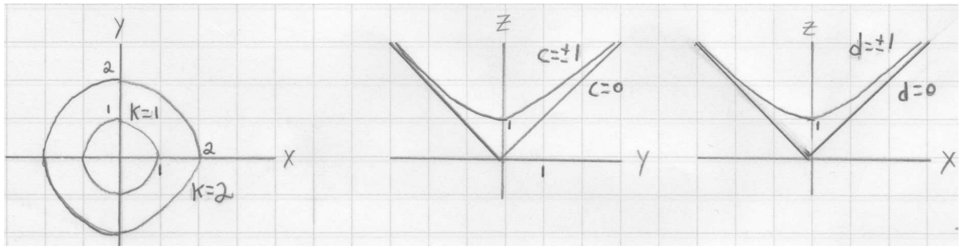
b) The domain of f is $16 - x^2 + y^2 \geq 0 \Rightarrow x^2 - y^2 \leq 16$. The range is $z \geq 0$.

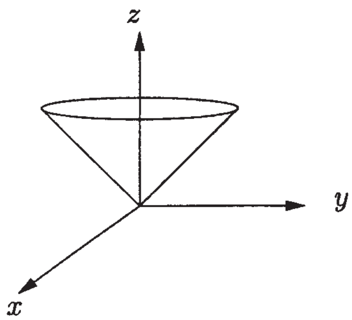


1.2 **Exercise 1:**



Exercise 2:





Answers to Chapter 2

2.3 **Exercise 1:** Show that $\lim_{y\to 0} f(0, y) = 0 \neq 1$. Thus $f(x, y)$ does not approach a unique value as $(x, y) \to (0, 0)$.

Exercise 2: Show that $\lim_{x\to 0} f(x, mx^3) = \dfrac{m}{1 + m^2}$. Thus $f(x, y)$ does not approach a unique value as $(x, y) \to (0, 0)$.

Exercise 3: Show that $\lim_{x\to 1} f(x, 0)$ does not exist. Thus $\lim_{(x,y)\to(1,0)} f(x, y)$ does not exist.

2.4 **Exercise 1:** If $m(x, y) \leq f((x, y)) \leq M(x, y)$ and $\lim_{(x,y)\to(a,b)} m(x, y) = L = \lim_{(x,y)\to(a,b)} M(x, y)$, then $\lim_{(x,y)\to(a,b)} f(x, y) = L$. To change this into our version take $m(x, y) = -B(x, y) + L$ and $M(x, y) = B(x, y) + L$.

Exercise 2: $|x^3 - y^3| \leq |x^3| + |y^3| \leq \big(|x| + |y|\big)(x^2 + y^2)$. Equality holds if and only if $x = 0$ or $y = 0$.

Exercise 3: Show that $\lim_{x\to 0} f(x, mx) = -1$. Hence the limit may exist and equal -1. A suitable inequality is

$$0 \leq |f(x, y) - L| = \left| \frac{x^2(x - 1) - y^2}{x^2 + y^2} - (-1) \right| = \frac{|x^3|}{x^2 + y^2} \leq |x|.$$

The Squeeze Theorem implies that $\lim_{(x,y)\to(0,0)} f(x, y) = L = -1$.

Answers to Chapter 3

3.1 **Exercise 1:** One example is $f(x) = \begin{cases} 5, & \text{if } x \geq 1 \\ 0, & \text{if } x < 1. \end{cases}$

Exercise 2: Use $\frac{|xy|}{|x|+|y|} \le \frac{|x|(|x|+|y|)}{|x|+|y|} = |x|$ to prove that $\lim\limits_{(x,y)\to(0,0)} f(x,y) = 0 = f(0,0)$. Hence f is continuous at $(0,0)$.

3.2 **Exercise 1:** By the limit theorem and the definition of product:

$$\begin{aligned}\lim_{(x,y)\to(a,b)} (fg)(x,y) &= \lim_{(x,y)\to(a,b)} f(x,y) \lim_{(x,y)\to(a,b)} g(x,y)) \\ &= f(a,b)g(a,b), \quad \text{by the hypothesis} \\ &= (fg)(a,b), \quad \text{by definition of product.}\end{aligned}$$

Therefore, by definition of continuity, fg is continuous at (a,b).

Exercise 2: Apply the limit theorems, the definition of quotient, and the definition of continuity as in exercise 1. $g(a,b) \neq 0$ is used explicitly when you use the definition of quotient and the limit theorem.

Note: Since $g(a,b) \neq 0$ and g is continuous at (a,b), $g(x,y) \neq 0$ for all (x,y) in some neighborhood of (a,b).

Exercise 3: For $f(x,y) = k$ we have $\lim\limits_{(x,y)\to(a,b)} k = k = f(a,b)$. For $f(x,y) = x$ we have $\lim\limits_{(x,y)\to(a,b)} x = a = f(a,b)$. For $f(x,y) = y$ we have $\lim\limits_{(x,y)\to(a,b)} y = b = f(a,b)$. So, they are all continuous on their domain.

Exercise 4: $h(x,y) = (xy)^\pi = e^{\pi \ln(xy)}$. Use the coordinate and constant functions, $e^{(\cdot)}$ and $\ln(\cdot)$. Use the product and composition theorems.

Exercise 5: Use the coordinate functions, the constant function, $|\cdot|$ and $\sin(\cdot)$. Use the sum, product, quotient and composition theorems.

Exercise 6: Show that $\lim\limits_{x\to 0} f(x,mx) = \dfrac{m}{1+m^2}$. Therefore $\lim\limits_{(x,y)\to(0,0)} f(x,y)$ does not exist, and you cannot make f continuous at $(0,0)$.

3.3 **Exercise 1:** By the Continuity Theorems $f(x,y) = \ln(1+e^{\sin xy})$ is continuous for all (x,y). Consequently $\lim\limits_{(x,y)\to(1,\pi)} f(x,y) = f(1,\pi) = \ln 2$.

Answers to Chapter 4

4.1 **Exercise 1:** $f_x = y^2\cos(xy^2)$, $f_y = 2xy\cos(xy^2)$.

Exercise 2: Show that for $a \neq 0$, $h \neq 0$,

$$\frac{f(a+h,-a)-f(a,-a)}{h} = \frac{(3a^2+3ah+h^2)^{1/3}}{h^{2/3}}.$$

Since $\lim\limits_{h\to 0} \dfrac{(3a^2+3ah+h^2)^{1/3}}{h^{2/3}} = +\infty$ for $a \neq 0$, $\dfrac{\partial f}{\partial x}(a,-a)$ does not exist.

Exercise 3: Show that $\dfrac{f(h,a)-f(0,a)}{h} = \dfrac{|h||a-1|}{h}$. Hence, with $a = 0$, $f_x(0,0)$ does not exist, but with $a = 1$, $f_x(0,1) = 0$.

Exercise 4: $\frac{\partial f}{\partial x}(a,b,c) = \lim_{h\to 0} \frac{f(a+h,b,c)-f(a,b,c)}{h}$ provided the limit exists. Similarly for $\frac{\partial f}{\partial y}$ and $\frac{\partial f}{\partial z}$.

4.2 **Exercise 1:** $f_{xx} = \dfrac{2(y^2-x^2)}{(x^2+y^2)^2}$; by symmetry $f_{yy} = \dfrac{2(x^2-y^2)}{(x^2+y^2)^2}$.

Exercise 2: $f_{xy} = x^{y-1}(1 + y\ln x) = f_{yx}$.

Exercise 3:

(a) We have

$$f_x(0,y) = \lim_{h\to 0} \frac{f(0+h,y)-f(0,0)}{h} = \lim_{h\to 0} \frac{hy\frac{h^2-y^2}{h^2+y^2}}{h} = -y$$

A similar computations shows that $f_y(x,0) = x$.

(b) We have

$$f_{xy}(0,0) = (f_x)_y(0,0) = \lim_{h\to 0} \frac{f_x(0,0+h)-f_x(0,0)}{h}$$

Using part (a), the above limit becomes

$$\lim_{h\to 0} \frac{-h-0}{h} = -1$$

The computation of $f_{yx}(0,0) = (f_y)_x(0,0)$ is similar.

(c) We can check directly that $f_{xy}(x,y)$ is not continuous at $(0,0)$. A straightforward computation shows that, for $(x,y) \neq (0,0)$, we have

$$f_x(x,y) = \frac{y(x^4+4x^2y^2-y^4)}{(x^2+y^2)^2}$$

and therefore

$$f_{xy}(x,y) = \frac{x^6+9x^4y^2-9x^2y^4-y^6}{(x^2+y^2)^3}$$

Thus

$$f_{xy}(x,y) = \begin{cases} \dfrac{x^6+9x^4y^2-9x^2y^4-y^6}{(x^2+y^2)^3} & \text{if } (x,y) \neq (0,0), \\ -1 & \text{if } (x,y) = (0,0) \end{cases}$$

We claim that $\lim\limits_{(x,y)\to(0,0)} f_{xy}(x,y)$ does not exist. Indeed, if we let (x,y) approach $(0,0)$ along the line $y=0$, we find that $f_{xy}(x,y) \to 1$, while if we approach along $x=0$, we find that $f_{xy}(x,y) \to -1$.

4.3 **Exercise 1:** $z = 5 + \frac{3}{5}(x-3) - \frac{4}{5}(y+4)$.

Exercise 2: The equation of the tangent plane at $(a, b, \sqrt{a^2+b^2})$ is

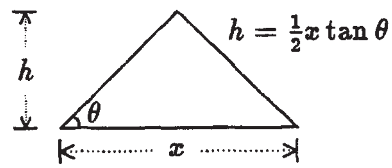
$$z = \sqrt{a^2+b^2} + \frac{a}{\sqrt{a^2+b^2}}(x-a) + \frac{b}{\sqrt{a^2+b^2}}(y-b).$$

Substituting $(x,y) = (0,0)$ gives $z = 0$.

4.4 **Exercise 2:** Let $f(x,y) = \sqrt{\sin x + \tan y}, \quad (a,b) = \left(0, \frac{\pi}{4}\right)$.

Show that $\sqrt{\sin x + \tan y} \approx 1 + \frac{1}{2}x + \left(y - \frac{\pi}{4}\right)$, for (x,y) sufficiently close to $\left(0, \frac{\pi}{4}\right)$. Hence $\sqrt{\sin\left(\frac{1}{10}\right) + \tan\left(\frac{3}{4}\right)} \approx 1.015$. [Calculator value 1.0156]

Exercise 4: The area A is $A = f(x, \theta) = \frac{1}{4}x^2 \tan\theta$. Show that $\Delta A \approx -0.48m^2$. [Calculator -0.4626]



4.5 **Exercise 1:** Let $f(x, y, z) = xyz$, $(a, b, c) = (5, 7, 10)$. Show that

$$xyz \approx 350 + 70(x - 5) + 50(y - 7) + 35(z - 10).$$

Hence $4.99 \times 7.01 \times 9.99 \approx 349.45$. [Calculator 349.4492]

Answers to Chapter 5

5.1 **Exercise 1:** Use the definition of partial derivative to show that $f_x(0, 0) = 1$ and $f_y(0, 0) = 0$. It follows that $\frac{|R_{1,(0,0)}(x,y)|}{\|(x,y)-(0,0)\|} = g(x, y)$, where $g(x, y) = \frac{|xy^2|}{(x^2+y^2)^{3/2}}$. Show that $\lim_{x\to 0} g(x, x) = 2^{-\frac{3}{2}} \neq 0$.

Exercise 2: Show that $\dfrac{|R_{1,(0,0)}(x, y)|}{\|(x, y) - (0, 0)\|} = \dfrac{|xy|}{\sqrt{x^2 + y^2}}$, and that $0 \leq \dfrac{|xy|}{\sqrt{x^2 + y^2}} \leq |y|$.

Exercise 3: Show that $\dfrac{f(h, 1) - f(0, 1)}{h} = \dfrac{|h|}{h}$, so that $f_x(0, 1)$ does not exist. Hence, f can not be differentiable.

Exercise 4: One possibility is $f(x, y) = \sqrt{(x - 1)^2 + (y - 2)^2}$.

5.2 **Exercise 1:** f is not differentiable at $\mathbf{a}$, since if it were, Theorem 1 would imply that f is continuous at $\mathbf{a}$, a contradiction.

Exercise 2: See example 1 in section 5.1.

5.3 **Exercise 1:** Show that $f_x(0, 0) = 0 = f_y(0, 0)$, and $\dfrac{|R_{1,(0,0)}(x, y)|}{\|(x, y) - (0, 0)\|} = (x^2 + y^2)^{1/6}$.

Exercise 2: Since the partial derivatives of f_x (that is, f_{xx} and f_{xy}) are continuous, f_x is differentiable and hence continuous. Similarly, the partial derivatives of f_y are continuous, hence f_y is differentiable and thus continuous. Therefore f is differentiable and hence also continuous.

5.4 **Exercise 1:** Use the approximation formula to obtain

$$g(x, y) = \sqrt{1 + 3\tan x + \sin y} \approx 2 + \frac{3}{2}\left(x - \frac{\pi}{4}\right) + \frac{1}{4}y.$$

Use the continuity theorems to prove that g has continuous partials. Then theorem 2 implies that the approximation is valid for (x, y) sufficiently close to $\left(\frac{\pi}{4}, 0\right)$.

Exercise 2:

(a) TRUE. If f is not continuous at $(0, 0)$, then f is not differentiable at $(0, 0)$ by Theorem 5.2.1. So, at least one of f_x and f_y must not be continuous at $(0, 0)$ as otherwise f would be differentiable at $(0, 0)$ by Theorem 5.3.2.

(b) FALSE. The function $f(x,y) = \begin{cases} \frac{x^2y}{x^2+y^2} & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$ is continuous at $(0,0)$ and both its partial derivatives exist, but it is not differentiable at $(0,0)$.

(c) FALSE. Use the same function as in (b).

(d) FALSE. Use the same function as in (b).

(e) TRUE. By definition of differentiability.

(f) FALSE. The function $f(x,y) = \begin{cases} (x^2+y^2)\sin\left(\frac{1}{x^2+y^2}\right) & \text{if } (x,y) \neq (0,0) \\ 0 & \text{if } (x,y) = (0,0) \end{cases}$ is differentiable at $(0,0)$, but both its partial derivatives are not continuous at $(0,0)$.

Answers to Chapter 6

6.1 **Exercise 1:** f is not differentiable at $(0,0)$.

Exercise 2: $\dfrac{dT}{dt}(0) = \frac{8}{5}$.

Exercise 3: $f'(1) = -2$. Assume that g is differentiable at $(2,0)$.

Exercise 4: $g'(t) = f_x(\cos t, \sin t)(-\sin t) + f_y(\cos t, \sin t)(\cos t)$; $g'\left(\frac{\pi}{3}\right) = \frac{1}{2}$.

Exercise 5: $g'(t) = \nabla F(t, t^2, t^3) \cdot (1, 2t, 3t^2)$; $g'(1) = 6$.

6.2 **Exercise 1:** Repeat what was done near the beginning of section 6.1 and use the linear approximation again to evaluate $\frac{\Delta x}{\Delta t}$ and $\frac{\Delta y}{\Delta t}$. We need f to be differentiable to ensure the linear approximation is a good approximation.

Exercise 2: $\frac{\partial g}{\partial y}(x,y) = 2xD_1f(2xy, x^2-y^2) - 2yD_2f(2xy, x^2-y^2)$, so $\frac{\partial g}{\partial y}(1,1) = -2$.

Exercise 3: $g'(t) = D_1f(h(t)+t, h(t)-t)(h'(t)+1) + D_2f(h(t)+t, h(t)-t)(h'(t)-1)$.

Exercise 4: $g'(1) = 2$.

Exercise 5: Repeat what was done near the beginning of section 6.1.

Exercise 6:

$$\frac{\partial g}{\partial x}(x,y) = (1)f(2xy, x^2-y^2) + x\Big[(2y)D_1f(2xy, x^2-y^2) + 2xD_2f(2xy, x^2-y^2)\Big].$$
$$\frac{\partial g}{\partial x}(1,1) = 9.$$

Exercise 7:

$$\frac{\partial u}{\partial s} = D_1f(\cdots)\frac{\partial x}{\partial s} + D_2f(\cdots)\frac{\partial y}{\partial s} + D_3f(\cdots)(1)$$

where $(\cdots) = \big(x(s,t), y(s,t), s, t\big)$.

6.3 **Exercise 1:** Assume that g has continuous second partials.

Exercise 2: $f_x = yg'(xy)$, $f_y = xg'(xy)$, $f_{xx} = y^2g''(xy)$, $f_{yy} = x^2g''(xy)$, so $x^2f_{xx} = x^2y^2g''(xy) = y^2g''(xy)$. Assume that g has a continuous second derivative.

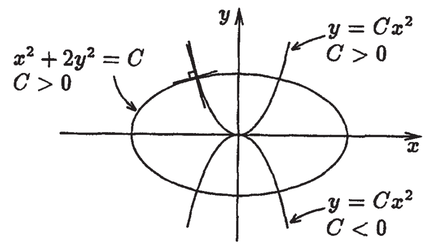
Answers to Chapter 7

7.1 **Exercise 1:** $D_{\hat{u}}f(1,-1,2) = \frac{4}{3e^2}$, with $\hat{u} = \left(\frac{1}{3}, \frac{2}{3}, -\frac{2}{3}\right)$.

7.2 **Exercise 1:** The largest rate of change is $\|\nabla f(0,1)\| = \sqrt{5}$, and occurs in the direction $(1,2)$.

Exercise 2: Give f and $\mathbf{a}$ such that $\nabla f(\mathbf{a}) = \mathbf{0}$, e.g. $f(x,y) = x^2 + y^2$, $\mathbf{a} = (\mathbf{0}, \mathbf{0})$. The tangent plane is horizontal at $\mathbf{a}$.

Exercise 3: Show that $\nabla f \cdot \nabla g = 0$, and apply Theorem 2.



7.3 **Exercise 1:** $(x-1) + 2(y-1) + 3\sqrt{3}(z - \sqrt{3}) = 0$.

Exercise 2: $8(x-1) - 3(y-2) + (z+2) = 0$.

Answers to Chapter 8

8.1 **Exercise 1:** $P_{2,\mathbf{a}}(x,y) = \frac{2}{3} - (x-1)^2 + \frac{1}{2}y^2$.

8.2 **Exercise 1:** We have $f_{xx} = 4e^{-2x+y}$, $f_{xy} = -2e^{-2x+y}$, and $f_{yy} = e^{-2x+y}$. Since $f \in C^2$, by Taylor's Theorem there is a point $\mathbf{c}$ such that

$$\begin{aligned}\left|R_{1,(1,1)}(x,y)\right| &= \frac{1}{2}\left|f_{xx}(\mathbf{c})(x-1)^2 + 2f_{xy}(\mathbf{c})(x-1)(y-1) + f_{yy}(\mathbf{c})(y-1)^2\right| \\ &\leq \frac{1}{2}\left[|f_{xx}(\mathbf{c})|\,(x-1)^2 + 2|f_{xy}(\mathbf{c})||(x-1)||(y-1)| + |f_{yy}(\mathbf{c})|(y-1)^2\right],\end{aligned}$$

by the triangle inequality. Thus, on $0 \leq x \leq 1$ and $0 \leq y \leq 1$ we have

$$|f_{xx}| \leq 4e, \qquad |f_{xy}| \leq 2e, \qquad |f_{yy}| \leq e.$$

Hence,

$$\begin{aligned}\left|R_{1,(1,1)}(x,y)\right| &\leq 2e(x-1)^2 + 2e|x-1||y-1| + \frac{1}{2}e(y-1)^2 \\ &\leq 2e(x-1)^2 + e(x-1)^2 + e(y-1)^2 + 2e(y-1)^2 \\ &= 3e[(x-1)^2 + (y-1)^2]\end{aligned}$$

8.3 **Exercise 1:**

$$\begin{aligned}P_{3,(a,b)} = P_{2,(a,b)}(x,y) &+ \frac{1}{6}f_{xxx}(a,b)(x-a)^3 + \frac{1}{2}f_{xxy}(x-a)^2(y-b) \\ &+ \frac{1}{2}f_{xyy}(x-a)(y-b)^2 + \frac{1}{6}f_{yyy}(a,b)(y-b)^3\end{aligned}$$

Answers to Chapter 9

9.1 **Exercise 1:** $f_x = y(1+x)e^{x-y}$, $f_y = x(1-y)e^{x-y}$; $(0,0)$ and $(-1,1)$.

Exercise 2: Critical points are $\left(0, \frac{\pi}{2} + k\pi\right)$, $k \in \mathbb{Z}$.

Exercise 3: One possibility is a linear function, e.g. $f(x, y) = 2x + 3y$.

9.2 **Exercise 2:** One critical point $(0,0)$, a saddle point.

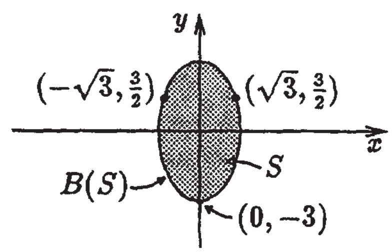
Exercise 3: The critical points are $(\pm 1, 0)$ and $\left(0, \pm\frac{1}{\sqrt{3}}\right)$. $Hf(\pm 1, 0) = \begin{bmatrix} 0 & \pm 2 \\ \pm 2 & 0 \end{bmatrix}$, indefinite; $(\pm 1, 0)$ are saddle points. $Hf\left(0, \frac{1}{\sqrt{3}}\right) = \begin{bmatrix} \frac{2\sqrt{3}}{3} & 0 \\ 0 & 2\sqrt{3} \end{bmatrix}$, positive definite; $\left(0, \frac{1}{\sqrt{3}}\right)$ is a local minimum point. Similarly, $\left(0, -\frac{1}{\sqrt{3}}\right)$ is a local maximum point.

Answers to Chapter 10

10.1 **Exercise 1:** 1. $I = [0, 2]$, $f(x) = \begin{cases} x, & \text{if } 0 \le x < 1 \\ x - 2 & \text{if } 1 \le x \le 2. \end{cases}$

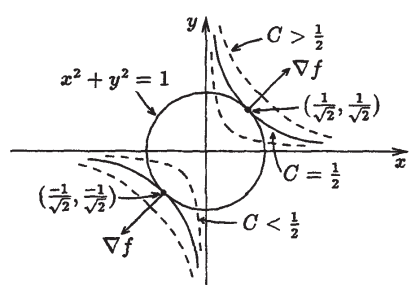
2. $I = (0, \frac{\pi}{2})$, $f(x) = \tan x$. 3. $I = [1, \infty)$, $f(x) = \frac{1}{x}$.

10.2 **Exercise 1:** Critical points of f are $(1,0)$ and $(-1,0)$. On the boundary, $g(t) = f(2\cos t, 3\sin t) = 3\sin t(3 - 4\sin^2 t)$. Critical points of g are $t = \frac{\pi}{6}, \frac{5\pi}{6}, \frac{7\pi}{6}, \frac{11\pi}{6}, \frac{\pi}{2}, \frac{3\pi}{2}$. Maximum value of f is 3, and occurs at $\left(\pm\sqrt{3}, \frac{3}{2}\right)$ and $(0, -3)$.



Exercise 2: Maximum value is $\frac{1}{4}$, and occurs at $\left(\frac{1}{2}, \frac{1}{2}\right)$.

10.3 **Exercise 1:** Maximum value is $\frac{1}{2}$ and occurs at $\left(\pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}}\right)$.



Exercise 2: Maximum value is $24 + 4\sqrt{6}$ at $(2\sqrt{6}, 0)$ and the minimum value is -1 at $(-1, 0)$.

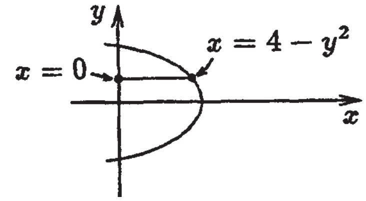
Exercise 3: The closest points are $(0, 0, \pm 1)$.

Answers to Chapter 11

11.1 **Exercise 1:** The integral equals the number of people in the region D.

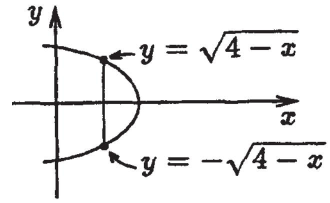
11.2 **Exercise 1:**

a) $D: 0 \le x \le 4 - y^2, -2 \le y \le 2.$



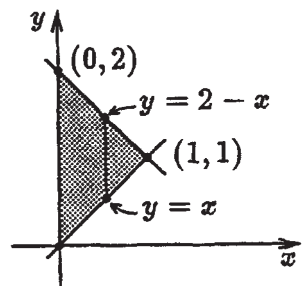
$$I = \iint_D (x+y)\, dA = \int_{-2}^{2} \int_{x=0}^{4-y^2} (x+y)\, dx\, dy = \frac{256}{15}.$$

b) $D: -\sqrt{4-x} \le y \le \sqrt{4-x}, 0 \le x \le 4.$



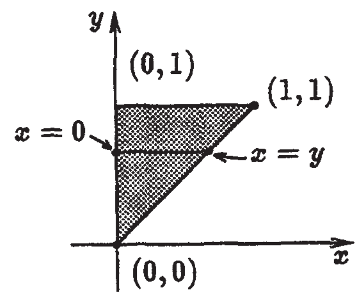
$$I = \int_0^4 \int_{-\sqrt{4-x}}^{\sqrt{4-x}} (x+y)\, dy\, dx.$$

Exercise 2: $D: x \le y \le 2 - x, 0 \le x \le 1.$



$$\iint_D y\, dA = \int_0^1 \int_x^{2-x} y\, dy\, dx = 1.$$

Exercise 3: $D: 0 \le x \le y, 0 \le y \le 1.$



$$\iint_D e^{-y^2}\, dA = \int_0^1 \int_0^y e^{-y^2}\, dx\, dy = \frac{e-1}{2e}.$$

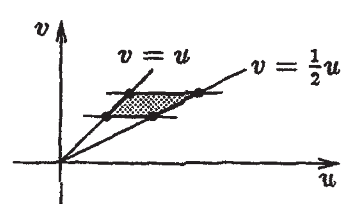
Exercise 4: $\Delta V \approx (4 - x^2 - y^2)\Delta A$; D is a rectangle.

$$V = \iint_D (4 - x^2 - y^2)\, dA = \int_0^1 \int_0^1 (4 - x^2 - y^2)\, dy\, dx = \frac{10}{3}.$$

Answers to Chapter 12

12.1 **Exercise 1:** The image is $\left(u - \frac{1}{2}\right)^2 + \left(v + \frac{1}{2}\right)^2 = \frac{1}{2}$.

Exercise 2: $F(S) = \{(u, v) \mid v \le u \le 2v, 2 \le v \le 3\}$.



12.2 **Exercise 1:** $\begin{bmatrix} \Delta u \\ \Delta v \end{bmatrix} \approx DF(1,0) \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}$, for $\Delta x, \Delta y$ sufficiently small. $F(0.95, 0.1) \approx (0.05, -0.15)$. [Calculator $(0.0488, -0.1625)$]

12.3 **Exercise 1:** a) $D(F \circ G) = \begin{bmatrix} \frac{4uvx}{\sqrt{2x^2+2y^2}} + 2u^2 & \frac{4uvy}{\sqrt{2x^2+2y^2}} + 2yu^2 \\ \frac{2xve^{uv-1}}{\sqrt{2x^2+2y^2}} + 2ue^{uv-1} & \frac{2yve^{uv-1}}{\sqrt{2x^2+2y^2}} + 2yue^{uv-1} \end{bmatrix}$

b) $D(G \circ F)(1,1) = DG(1,1)DF(1,1) = \begin{bmatrix} 3 & 2 \\ 6 & 4 \end{bmatrix}$.

c) $(G \circ F)(1.01, 0.98) \approx (G \circ F)(1,1) + D(G \circ F)(1,1)\begin{bmatrix} .01 \\ -0.02 \end{bmatrix} = \begin{bmatrix} 1.99 \\ 2.98 \end{bmatrix}$.

Answers to Chapter 13

13.1 **Exercise 1:** $\dfrac{\partial(x,y)}{\partial(r,\theta)} = \det\begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix} = r$.

Exercise 2: This involves solving a quadratic equation. In order to choose the appropriate sign, ensure that the image of $(x,y) = (1,-2)$, i.e. the point $(u,v) = (-3,-1)$, is mapped by F^{-1} onto $(1,-2)$ again.

13.2 **Exercise 1:** $\left|\frac{\partial(u,v)}{\partial(x,y)}\right| = x^2y$, so if $\frac{1}{2} \le x \le 1$ and $\frac{1}{2} \le y \le 1$, then $x^2y \le 1$. Thus the image of S under F will have less area.

Exercise 2: Show that $\left|\dfrac{\partial(u,v)}{\partial(x,y)}\right| = 1$.

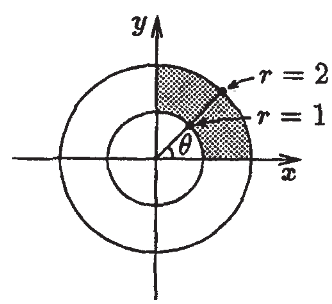
Exercise 3: Show that $\dfrac{\partial(u,v)}{\partial(x,y)} = 1$.

13.3 **Exercise 1:** Observe that $3x^2 + 2xy + y^2 = 2x^2 + (x+y)^2$, so take $u = \sqrt{2}x$ and $v = x + y$.

Exercise 2: Let $u = xy$, $v = xz$, and $w = yz$. The cube is $1 \le u \le 3$, $1 \le v \le 3$, $2 \le w \le 4$.

13.4 **Exercise 3:** Use polar coordinates.

$$\iint_{D_{xy}} \frac{1}{\sqrt{x^2+y^2}}\, dx\, dy = \int_0^{\frac{\pi}{2}} \int_1^2 \frac{1}{r}(r)\, dr\, d\theta = \frac{\pi}{2}.$$



Exercise 4: Let $(u,v) = \left(xy, \frac{y}{x}\right)$. Use the inverse property of the Jacobian to show that $\dfrac{\partial(x,y)}{\partial(u,v)} = \frac{1}{2v}$. Then

$$I = \int_2^3 \int_1^e u\left(\frac{1}{2v}\right) dv\, du = \frac{5}{4}.$$

Answers to Chapter 14

14.2 **Exercise 2:** $$\iiint_D z\,dV = \int_0^c \int_0^{a\left(1-\frac{z}{c}\right)} \int_0^{b\left(1-\frac{x}{a}-\frac{z}{c}\right)} z\,dy\,dx\,dz.$$

Exercise 3: A triple integral can be written as an iterated integral in $3! = 6$ ways.

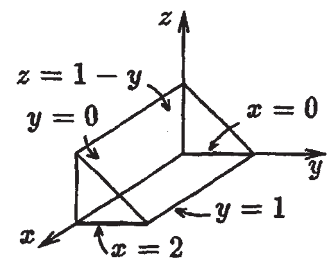
Exercise 4: Refer to the diagram in Example 2. The iterated integral is

$$\int_0^2 \int_{2-y}^{6-2y} \int_0^{\sqrt{4-y^2}} \frac{z}{4-y}\,dz\,dx\,dy.$$

In order to integrate first with respect to y, you would have to decompose D into several pieces.

Exercise 5: D is described by the inequalities $0 \le z \le 1-y$, $0 \le y \le 1$, $0 \le x \le 2$.

$$\iiint_D y\,dV = \int_0^2 \int_0^1 \int_0^{1-y} y\,dz\,dy\,dx = \frac{1}{3}.$$



14.3 **Exercise 2:** The solid is described by the inequalities $1 \le x+y \le 2$, $-1 \le x-y \le 1$, $0 \le x+y+z \le 3$. Let $(u,v,w) = (x+y, x-y, x+y+z)$. Show that

$$\left|\frac{\partial(x,y,z)}{\partial(u,v,w)}\right| = \frac{1}{2}.$$

Then

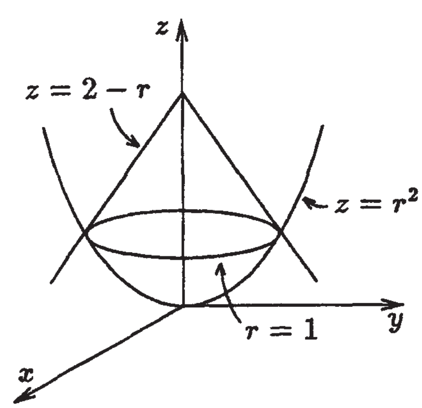
$$V = \int_1^2 \int_{-1}^1 \int_0^3 1\left(\frac{1}{2}\right) dw\,dv\,du = 3.$$

Exercise 3: Use cylindrical coordinates.

$$M = \int_0^{2\pi} \int_0^1 \int_0^2 \frac{k(2-z)}{1+r^2}(r)\,dz\,dr\,d\theta = 2\pi k \ln 2.$$

Exercise 4: Use cylindrical coordinates. The paraboloid is $z = r^2$, and the lower part of the cone is $z = 2-r$, and they intersect in the circle $r = 1$.

$$V = \int_0^{2\pi} \int_0^1 \int_{r^2}^{2-r} 1(r)\,dz\,dr\,d\theta = \frac{5\pi}{6}.$$



Exercise 6: Use the hint, and show that $\dfrac{\partial(x,y,z)}{\partial(u,v,w)} = abc$.

$$V = \iiint\limits_{D_{uvw}} abc\,du\,dv\,dw,$$

where D_{uvw} is given by $u^2 + v^2 + w^2 = 1$. Replace u, v, w by spherical coordinates and show that

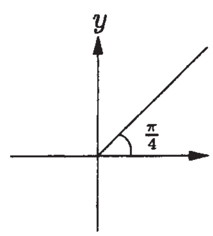
$$V = \frac{4}{3}\pi abc.$$

Exercise 7: Use spherical coordinates, and refer to the diagram in the text.

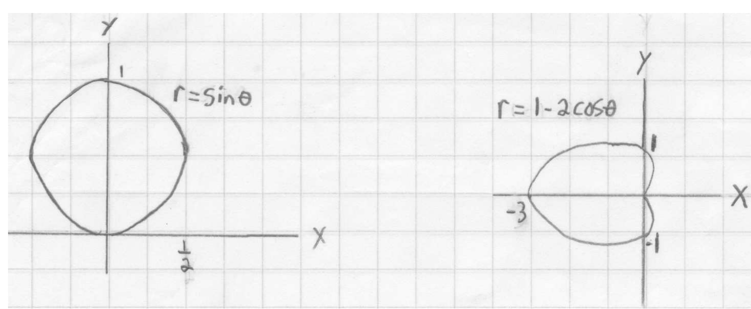
$$V = \int_0^{2\pi}\int_0^{\alpha}\int_0^{b} r^2 \sin\phi\,dr\,d\phi\,d\theta = \frac{2}{3}\pi b^3(1-\cos\alpha)$$

Answers to Appendix B

B.1 **Exercise 1:**



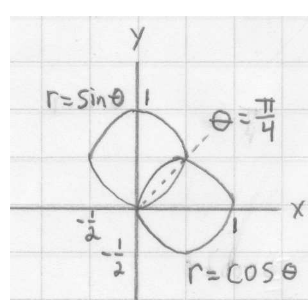
Exercise 2:



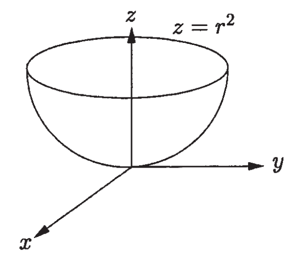
Exercise 3: $r^2\cos 2\theta = 1$.

Exercise 4: $A = 2\int_0^{\pi/2} \frac{1}{2}[2\sqrt{\sin 2\theta}]^2\,d\theta = 4$.

Exercise 5: $A = \int_0^{\pi/4} \frac{1}{2}\sin^2\theta\,d\theta + \int_{\pi/4}^{\pi/2} \frac{1}{2}\cos^2\theta\,d\theta = \frac{\pi}{8} - \frac{1}{4}$.



B.2 **Exercise 1:**



Exercise 2: $z = \sin\theta$, $r \neq 0$.

B.3 **Exercise 1:** $\rho = 2\cos\theta\sin\phi$, $-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}$.

Index