THE UNIVERSITY of WISCONSIN MADISON

MATH 222

Second Semester Calculus

Fall 2013

Typeset:December 26, 2013

Math 222 – 2nd Semester Calculus
Lecture notes version 1.0 (Fall 2013)

This is a self contained set of lecture notes for Math 222. The notes were written by Sigurd Angenent, as part ot the MIU calculus project. Some problems were contributed by A.Miller.

The LaTeX files, as well as the Inkscape and Octave files which were used to produce these notes are available at the following web site

http://www.math.wisc.edu/~angenent/MIU-calculus

They are meant to be freely available for non-commercial use, in the sense that "free software" is free. More precisely:

Contents

Chapter 1. Methods of Integration 9
1. Definite and indefinite integrals 9
1.1. Definition 9
2. Problems 11
3. First trick: using the double angle formulas 11
3.1. The double angle formulas 11
4. Problems 12
5. Integration by Parts 14
5.1. The product rule and integration by parts 14
5.2. An Example – Integrating by parts once 14
5.3. Another example 14
5.4. Example – Repeated Integration by Parts 15
5.5. Another example of repeated integration by parts 15
5.6. Example – sometimes the factor $G'(x)$ is invisible 15
5.7. An example where we get the original integral back 16
6. Reduction Formulas 16
6.1. First example of a reduction formula 17
6.2. Reduction formula requiring two partial integrations 18
6.3. A reduction formula where you have to solve for I_n 18
6.4. A reduction formula that will come in handy later 19
7. Problems 20
8. Partial Fraction Expansion 21
8.1. Reduce to a proper rational function 21
8.2. Example 22
8.3. Partial Fraction Expansion: The Easy Case 22
8.4. Previous Example continued 23
8.5. Partial Fraction Expansion: The General Case 24
8.6. Example 24
8.7. A complicated example 25
8.8. After the partial fraction decomposition 25
9. Problems 26
10. Substitutions for integrals containing the expression $\sqrt{ax^2+bx+c}$ 27
10.1. Integrals involving $\sqrt{ax+b}$ 28
10.2. Integrals containing $\sqrt{1-x^2}$ 29
10.3. Integrals containing $\sqrt{a^2-x^2}$ 30
10.4. Integrals containing $\sqrt{x^2-a^2}$ or $\sqrt{a^2+x^2}$ 31
11. Rational substitution for integrals containing $\sqrt{x^2-a^2}$ or $\sqrt{a^2+x^2}$ 32
11.1. The functions $U(t)$ and $V(t)$ 32
12. Simplifying $\sqrt{ax^2+bx+c}$ by completing the square 35
12.1. Example 35
12.2. Example 36
12.3. Example 36
13. Problems 37
14. Chapter summary 37
15. Mixed Integration Problems 38

Chapter 2. Proper and Improper Integrals 41

1. Typical examples of improper integrals 41
1.1. Integral on an unbounded interval 41
1.2. Second example on an unbounded interval 42
1.3. An improper integral on a finite interval 43
1.4. A doubly improper integral 43
1.5. Another doubly improper integral 44
2. Summary: how to compute an improper integral 44
2.1. How to compute an improper integral on an unbounded interval 44
2.2. How to compute an improper integral of an unbounded function 44
2.3. Doubly improper integrals 45
3. More examples 45
3.1. Area under an exponential 45
3.2. Improper integrals involving x^{-p} 46
4. Problems 47
5. Estimating improper integrals 48
5.1. Improper integrals of positive functions 49
5.2. Comparison Theorem for Improper Integrals 50
5.3. The Tail Theorem 52
6. Problems 54

Chapter 3. First order differential Equations 57
1. What is a Differential Equation? 57
2. Two basic examples 57
2.1. Equations where the RHS does not contain y 57
2.2. The exponential growth example 58
2.3. Summary 58
3. First Order Separable Equations 59
3.1. Solution method for separable equations 59
3.2. A snag: 60
3.3. Example 60
3.4. Example: the snag in action 60
4. Problems 60
5. First Order Linear Equations 61
5.1. The Integrating Factor 61
5.2. An example 62
6. Problems 63
7. Direction Fields 64
8. Euler's method 65
8.1. The idea behind the method 65
8.2. Setting up the computation 66
9. Problems 67
10. Applications of Differential Equations 67
10.1. Example: carbon dating 67
10.2. Example: dating a leaky bucket 68
10.3. Heat transfer 69
10.4. Mixing problems 70
11. Problems 71

Chapter 4. Taylor's Formula 73
1. Taylor Polynomials 73
1.1. Definition 73
1.2. Theorem 73
2. Examples 74
2.1. Taylor polynomials of order zero and one 74
2.2. Example: Compute the Taylor polynomials of degree 0, 1 and 2 of $f(x) = e^x$ at $a = 0$, and plot them 75
2.3. Example: Find the Taylor polynomials of $f(x) = \sin x$ 76

2.4. Example: compute the Taylor polynomials of degree two and three of $f(x) = 1 + x + x^2 + x^3$ at $a = 3$ 77
3. Some special Taylor polynomials 78
4. Problems 79
5. The Remainder Term 80
5.1. Definition 80
5.2. Example 80
5.3. An unusual example, in which there *is* a simple formula for $R_n f(x)$ 81
5.4. Another unusual, but important example where we can compute $R_n f(x)$ 81
6. Lagrange's Formula for the Remainder Term 82
6.1. Theorem 82
6.2. Estimate of remainder term 82
6.3. How to compute e in a few decimal places 82
6.4. Error in the approximation $\sin x \approx x$ 83
7. Problems 84
8. The limit as $x \to 0$, keeping n fixed 84
8.1. Little-oh 84
8.2. Theorem 84
8.3. Definition 85
8.4. Example: prove one of these little-oh rules 85
8.5. Can we see that $x^3 = o(x^2)$ by looking at the graphs of these functions? 85
8.6. Example: Little-oh arithmetic is a little funny 86
8.7. Computations with Taylor polynomials 87
8.8. Theorem 87
8.9. How *NOT* to compute the Taylor polynomial of degree 12 of $f(x) = 1/(1 + x^2)$ 87
8.10. A much easier approach to finding the Taylor polynomial of **any** degree of $f(x) = 1/(1 + x^2)$ 88
8.11. Example of multiplication of Taylor polynomials 88
8.12. Taylor's formula and Fibonacci numbers 89
8.13. More about the Fibonacci numbers 90
9. Problems 91
10. Differentiating and Integrating Taylor polynomials 92
10.1. Theorem 92
10.2. Example: Taylor polynomial of $(1 - x)^{-2}$ 93
10.3. Example: Taylor polynomials of $\arctan x$ 93
11. Problems on Integrals and Taylor Expansions 93
12. Proof of Theorem 8.8 94
12.1. Lemma 94
13. Proof of Lagrange's formula for the remainder 95

Chapter 5. Sequences and Series 97
1. Introduction 97
1.1. A different point of view on Taylor expansions 97
1.2. Some sums with infinitely many terms 97
2. Sequences 99
2.1. Examples of sequences 99
2.2. Definition 99
2.3. Example: $\lim_{n\to\infty} \frac{1}{n} = 0$ 99
2.4. Example: $\lim_{n\to\infty} a^n = 0$ if $|a| < 1$ 100
2.5. Theorem 100
2.6. Sandwich theorem 100
2.7. Theorem 100
2.8. Example 101
2.9. Limits of rational functions 101
2.10. Example. Application of the Sandwich theorem. 101
2.11. Example: factorial beats any exponential 101
3. Problems on Limits of Sequences 102
4. Series 102

4.1. Definitions 102
4.2. Examples 103
4.3. Properties of series 104
4.4. Theorem 104
5. Convergence of Taylor Series 105
5.1. The GEOMETRIC SERIES converges for $-1 < x < 1$ 105
5.2. Convergence of the exponential Taylor series 106
5.3. The day that all Chemistry stood still 106
6. Problems on Convergence of Taylor Series 107
7. Leibniz' formulas for $\ln 2$ and $\pi/4$ 108
8. Problems 109

Chapter 6. Vectors 111
1. Introduction to vectors 111
1.1. Definition 111
1.2. Basic arithmetic of vectors 111
1.3. Some GOOD examples. 112
1.4. Two very, very BAD examples. 112
1.5. Algebraic properties of vector addition and multiplication 112
1.6. Example 113
2. Geometric description of vectors 113
2.1. Definition 114
2.2. Example 115
2.3. Example 115
2.4. Geometric interpretation of vector addition and multiplication 115
2.5. Example 115
3. Parametric equations for lines and planes 116
3.1. Example 117
3.2. Midpoint of a line segment. 117
4. Vector Bases 118
4.1. The Standard Basis Vectors 118
4.2. A Basis of Vectors (in general)* 118
4.3. Definition 118
4.4. Definition 118
5. Dot Product 119
5.1. Definition 119
5.2. Algebraic properties of the dot product 119
5.3. Example 119
5.4. The diagonals of a parallelogram 119
5.5. Theorem 120
5.6. The dot product and the angle between two vectors 120
5.7. Theorem 120
5.8. Orthogonal projection of one vector onto another 121
5.9. Defining equations of lines 122
5.10. Line through one point and perpendicular to another line 122
5.11. Distance to a line 123
5.12. Defining equation of a plane 124
5.13. Example 125
5.14. Example continued 125
5.15. Where does the line through the points $B(2,0,0)$ and $C(0,1,2)$ intersect the plane $\mathcal{P}$ from example 5.13? 126
6. Cross Product 127
6.1. Algebraic definition of the cross product 127
6.2. Definition 127
6.3. Example 127
6.4. Example 127
6.5. Algebraic properties of the cross product 127
6.6. Ways to compute the cross product 128

6.7. Example 128
6.8. The triple product and determinants 128
6.9. Definition 128
6.10. Theorem 129
6.11. Geometric description of the cross product 129
6.12. Theorem 130
6.13. Theorem 130
7. A few applications of the cross product 131
7.1. Area of a parallelogram 131
7.2. Example 131
7.3. Finding the normal to a plane 132
7.4. Example 132
7.5. Volume of a parallelepiped 133
8. Notation 133
Common abuses of notation that should be avoided 134
9. Problems–Computing and drawing vectors 135
10. Problems–Parametric equations for a line 137
11. Problems–Orthogonal decomposition of one vector with respect to another 137
12. Problems–The dot product 138
13. Problems–The cross product 138

$f(x) = \dfrac{dF(x)}{dx}$	$\displaystyle\int f(x)\,dx = F(x) + C$	
$(n+1)x^n = \dfrac{dx^{n+1}}{dx}$	$\displaystyle\int x^n\,dx = \frac{x^{n+1}}{n+1} + C$	$n \neq -1$
$\dfrac{1}{x} = \dfrac{d\ln\lvert x\rvert}{dx}$	$\displaystyle\int \frac{1}{x}\,dx = \ln\lvert x\rvert + C$	absolute values!!
$e^x = \dfrac{de^x}{dx}$	$\displaystyle\int e^x\,dx = e^x + C$	
$-\sin x = \dfrac{d\cos x}{dx}$	$\displaystyle\int \sin x\,dx = -\cos x + C$	
$\cos x = \dfrac{d\sin x}{dx}$	$\displaystyle\int \cos x\,dx = \sin x + C$	
$\tan x = -\dfrac{d\ln\lvert\cos x\rvert}{dx}$	$\displaystyle\int \tan x = -\ln\lvert\cos x\rvert + C$	absolute values!!
$\dfrac{1}{1+x^2} = \dfrac{d\arctan x}{dx}$	$\displaystyle\int \frac{1}{1+x^2}\,dx = \arctan x + C$	
$\dfrac{1}{\sqrt{1-x^2}} = \dfrac{d\arcsin x}{dx}$	$\displaystyle\int \frac{1}{\sqrt{1-x^2}}\,dx = \arcsin x + C$	
$f(x) + g(x) = \dfrac{dF(x) + G(x)}{dx}$	$\displaystyle\int \{f(x) + g(x)\}\,dx = F(x) + G(x) + C$	
$cf(x) = \dfrac{d\,cF(x)}{dx}$	$\displaystyle\int cf(x)\,dx = cF(x) + C$	
$F\dfrac{dG}{dx} = \dfrac{dFG}{dx} - \dfrac{dF}{dx}G$	$\displaystyle\int FG'\,dx = FG - \int F'G\,dx$	

To find derivatives and integrals involving a^x instead of e^x use $a = e^{\ln a}$, and thus $a^x = e^{x\ln a}$, to rewrite all exponentials as $e^{\cdots}$.

The following integral is also useful, but not as important as the ones above:

$$\int \frac{dx}{\cos x} = \frac{1}{2}\ln\frac{1+\sin x}{1-\sin x} + C \text{ for } \cos x \neq 0.$$

Table 1. The list of the standard integrals everyone should know

CHAPTER 1

Methods of Integration

The basic question that this chapter addresses is how to compute integrals, i.e.

Given a function $y = f(x)$ how do we find
a function $y = F(x)$ whose derivative is $F'(x) = f(x)$?

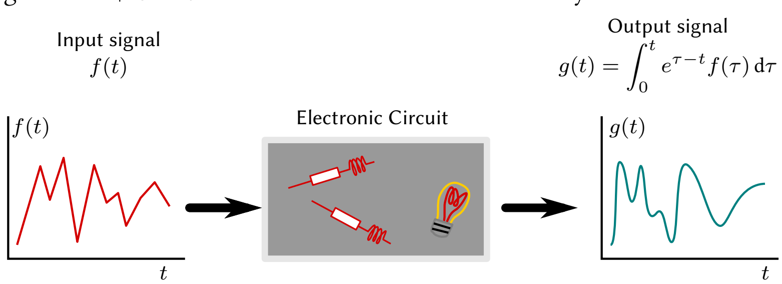
The simplest solution to this problem is to look it up on the Internet. Any integral that we compute in this chapter can be found by typing it into the following web page:

http://integrals.wolfram.com

Other similar websites exist, and more extensive software packages are available.

It is therefore natural to ask *why should we learn how to do these integrals?* The question has at least two answers.

First, there are certain *basic integrals* that show up frequently and that are relatively easy to do (once we know the trick), but that are not included in a first semester calculus course for lack of time. Knowing these integrals is useful in the same way that knowing things like "$2 + 3 = 5$" saves us from a lot of unnecessary calculator use.



The second reason is that we often are not really interested in specific integrals, but in general facts about integrals. For example, the output $g(t)$ of an electric circuit (or mechanical system, or a biochemical system, etc.) is often given by some integral involving the input $f(t)$. The methods of integration that we will see in this chapter give us the tools we need to understand why some integral gives the right answer to a given electric circuits problem, no matter what the input $f(t)$ is.

1. Definite and indefinite integrals

We recall some facts about integration from first semester calculus.

1.1. Definition. *A function $y = F(x)$ is called an* ***antiderivative*** *of another function $y = f(x)$ if $F'(x) = f(x)$ for all x.*

For instance, $F(x) = \frac{1}{2}x^2$ is an antiderivative of $f(x) = x$, and so is $G(x) = \frac{1}{2}x^2 + 2012$.

The ***Fundamental Theorem of Calculus*** states that if a function $y = f(x)$ is continuous on an interval $a \le x \le b$, then there always exists an antiderivative $F(x)$ of f,

Indefinite integral	Definite integral
$\int f(x)dx$ is a function of x.	$\int_a^b f(x)dx$ is a number.
By definition $\int f(x)dx$ is *any function $F(x)$ whose derivative is $f(x)$.*	$\int_a^b f(x)dx$ was defined in terms of Riemann sums and can be interpreted as "area under the graph of $y = f(x)$" when $f(x) \geq 0$.
If $F(x)$ is an antiderivative of $f(x)$, then so is $F(x) + C$. Therefore $\int f(x)\mathrm{d}x = F(x) + C$; an indefinite integral contains a constant ("$+C$").	$\int_a^b f(x)dx$ is one uniquely defined number; an indefinite integral does not contain an arbitrary constant.
x is not a dummy variable, for example, $\int 2xdx = x^2 + C$ and $\int 2tdt = t^2 + C$ are functions of different variables, so they are not equal. (See Problem 2.1.)	x is a dummy variable, for example, $\int_0^1 2xdx = 1,$ and $\int_0^1 2tdt = 1,$ so $\int_0^1 2xdx = \int_0^1 2tdt.$ Whether we use x or t the integral makes no difference.

Table 1. Important differences between definite and indefinite integrals

and one has

$$\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a). \tag{1}$$

For example, if $f(x) = x$, then $F(x) = \frac{1}{2}x^2$ is an antiderivative for $f(x)$, and thus $\int_a^b x\,\mathrm{d}x = F(b) - F(a) = \frac{1}{2}b^2 - \frac{1}{2}a^2$.

The best way of computing an integral is often to find an antiderivative F of the given function f, and then to use the Fundamental Theorem (1). *How to go about finding an antiderivative F for some given function f is the subject of this chapter.*

The following notation is commonly used for antiderivatives:

$$F(x) = \int f(x)\mathrm{d}x. \tag{2}$$

The integral that appears here does not have the integration bounds a and b. It is called an ***indefinite integral***, as opposed to the integral in (1) which is called a ***definite integral***. It is important to distinguish between the two kinds of integrals. Table 1 lists the main differences.

2. Problems

1. Compute the following integrals:

(a) $A = \int x^{-2}\,\mathrm{d}x$,

(b) $B = \int t^{-2}\,\mathrm{d}t$,

(c) $C = \int x^{-2}\,\mathrm{d}t$,

(d) $I = \int xt\,\mathrm{d}t$,

(e) $J = \int xt\,\mathrm{d}x$.

2. One of the following three integrals is not the same as the other two:

$$A = \int_1^4 x^{-2}\,\mathrm{d}x,$$

$$B = \int_1^4 t^{-2}\,\mathrm{d}t,$$

$$C = \int_1^4 x^{-2}\,\mathrm{d}t.$$

Which one? Explain your answer.

3. Which of the following inequalities are true?

(a) $\displaystyle\int_2^4 (1-x^2)\mathrm{d}x > 0$

(b) $\displaystyle\int_2^4 (1-x^2)\mathrm{d}t > 0$

(c) $\displaystyle\int (1+x^2)\mathrm{d}x > 0$

4. One of the following statements is correct. Which one, and why?

(a) $\displaystyle\int_0^x 2t^2\mathrm{d}t = \tfrac{2}{3}x^3$.

(b) $\displaystyle\int 2t^2\mathrm{d}t = \tfrac{2}{3}x^3$.

(c) $\displaystyle\int 2t^2\mathrm{d}t = \tfrac{2}{3}x^3 + C$.

3. First trick: using the double angle formulas

The first method of integration we see in this chapter uses trigonometric identities to rewrite functions in a form that is easier to integrate. This particular trick is useful in certain integrals involving trigonometric functions and while these integrals show up frequently, the "double angle trick" is not a general method for integration.

3.1. The double angle formulas.

The simplest of the trigonometric identities are the double angle formulas. These can be used to simplify integrals containing either $\sin^2 x$ or $\cos^2 x$.

Recall that

$$\cos^2\alpha - \sin^2\alpha = \cos 2\alpha \quad \text{and} \quad \cos^2\alpha + \sin^2\alpha = 1,$$

Adding these two equations gives

$$\cos^2\alpha = \frac{1}{2}\left(\cos 2\alpha + 1\right)$$

while subtracting them gives

$$\sin^2\alpha = \frac{1}{2}\left(1 - \cos 2\alpha\right).$$

These are the two double angle formulas that we will use.

3.1.1. *Example.* The following integral shows up in many contexts, so it is worth knowing:

$$\begin{aligned}\int \cos^2 x\,\mathrm{d}x &= \frac{1}{2}\int (1+\cos 2x)\mathrm{d}x \\ &= \frac{1}{2}\left\{x + \frac{1}{2}\sin 2x\right\} + C \\ &= \frac{x}{2} + \frac{1}{4}\sin 2x + C.\end{aligned}$$

Since $\sin 2x = 2\sin x \cos x$ this result can also be written as

$$\int \cos^2 x \,\mathrm{d}x = \frac{x}{2} + \frac{1}{2}\sin x \cos x + C.$$

3.1.2. *A more complicated example.* If we need to find

$$I = \int \cos^4 x \,\mathrm{d}x$$

then we can use the double angle trick once to rewrite $\cos^2 x$ as $\frac{1}{2}(1+\cos 2x)$, which results in

$$I = \int \cos^4 x \,\mathrm{d}x = \int \left\{\frac{1}{2}(1+\cos 2x)\right\}^2 \mathrm{d}x = \frac{1}{4}\int \left(1 + 2\cos 2x + \cos^2 2x\right) \mathrm{d}x.$$

The first two terms are easily integrated, and now that we know the double angle trick we also can do the third term. We find

$$\int \cos^2 2x \,\mathrm{d}x = \frac{1}{2}\int \left(1+\cos 4x\right) \mathrm{d}x = \frac{x}{2} + \frac{1}{8}\sin 4x + C.$$

Going back to the integral I we get

$$\begin{aligned} I &= \frac{1}{4}\int \left(1 + 2\cos 2x + \cos^2 2x\right) \mathrm{d}x \\ &= \frac{x}{4} + \frac{1}{4}\sin 2x + \frac{1}{2}\left(\frac{x}{4} + \frac{1}{8}\sin 4x\right) + C \\ &= \frac{3x}{8} + \frac{1}{4}\sin 2x + \frac{1}{16}\sin 4x + C \end{aligned}$$

3.1.3. *Example without the double angle trick.* The integral

$$J = \int \cos^3 x \,\mathrm{d}x$$

looks very much like the two previous examples, but there is very different trick that will give us the answer. Namely, substitute $u = \sin x$. Then $\mathrm{d}u = \cos x \mathrm{d}x$, and $\cos^2 x = 1 - \sin^2 x = 1 - u^2$, so that

$$\begin{aligned} J &= \int \cos^2 x \cos x \,\mathrm{d}x \\ &= \int (1 - \sin^2 x)\cos x \,\mathrm{d}x \\ &= \int (1 - u^2) \,\mathrm{d}u \\ &= u - \frac{1}{3}u^3 + C \\ &= \sin x - \frac{1}{3}\sin^3 x + C. \end{aligned}$$

In summary, the double angle formulas are useful for certain integrals involving powers of $\sin(\cdots)$ or $\cos(\cdots)$, but not all. In addition to the double angle identities there are other trigonometric identities that can be used to find certain integrals. See the exercises.

4. Problems

Compute the following integrals using the double angle formulas if necessary:

1. $\displaystyle\int (1+\sin 2\theta)^2 \, \mathrm{d}\theta$.

2. $\displaystyle\int (\cos\theta+\sin\theta)^2 \, \mathrm{d}\theta$.

3. Find $\displaystyle\int \sin^2 x \cos^2 x \, \mathrm{d}x$

(hint: use the other double angle formula $\sin 2\alpha = 2\sin\alpha\cos\alpha$.)

4. $\displaystyle\int \cos^5\theta \, \mathrm{d}\theta$

5. Find $\displaystyle\int \left(\sin^2\theta+\cos^2\theta\right)^2 \, \mathrm{d}\theta$

The double angle formulas are special cases of the following trig identities:

$$2\sin A\sin B = \cos(A-B) - \cos(A+B)$$
$$2\cos A\cos B = \cos(A-B) + \cos(A+B)$$
$$2\sin A\cos B = \sin(A+B) + \sin(A-B)$$

Use these identities to compute the following integrals.

6. $\displaystyle\int \sin x \sin 2x \ \ \mathrm{d}x$

7. $\displaystyle\int_0^\pi \sin 3x \sin 2x \ \ \mathrm{d}x$

8. $\displaystyle\int \left(\sin 2\theta - \cos 3\theta\right)^2 \, \mathrm{d}\theta$.

9. $\displaystyle\int_0^{\pi/2} \left(\sin 2\theta + \sin 4\theta\right)^2 \, \mathrm{d}\theta$.

10. $\displaystyle\int_0^\pi \sin kx \sin mx \, \mathrm{d}x$ where k and m are constant positive integers. Simplify your answer! (careful: after working out your solution, check if you didn't divide by zero anywhere.)

11. Let a be a positive constant and

$$I_a = \int_0^{\pi/2} \sin(a\theta)\cos(\theta)\,\mathrm{d}\theta.$$

(a) Find I_a if $a \neq 1$.

(b) Find I_a if $a = 1$. (Don't divide by zero.)

12. The input signal for a given electronic circuit is a function of time $V_{\text{in}}(t)$. The output signal is given by

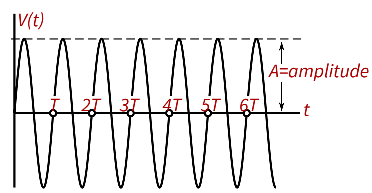
$$V_{\text{out}}(t) = \int_0^t \sin(t-s)V_{\text{in}}(s)\,\mathrm{d}s.$$

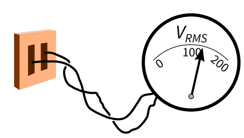
Find $V_{\text{out}}(t)$ if $V_{\text{in}}(t) = \sin(at)$ where $a > 0$ is some constant.

13. The alternating electric voltage coming out of a socket in any American living room is said to be 110Volts and 50Herz (or 60, depending on where you are). This means that the voltage is a function of time of the form

$$V(t) = A\sin(2\pi\frac{t}{T})$$

where $T = \frac{1}{50}$sec is how long one oscillation takes (if the frequency is 50 Herz, then there are 50 oscillations per second), and A is the *amplitude* (the largest voltage during any oscillation).





The 110 Volts that is specified is not the amplitude A of the oscillation, but instead it refers to the "*Root Mean Square*" of the voltage. By definition the R.M.S. of the oscillating voltage $V(t)$ is

$$110 = \sqrt{\frac{1}{T}\int_0^T V(t)^2 \mathrm{d}t}.$$

(it is the square root of the mean of the square of $V(t)$).

Compute the amplitude A.

5. Integration by Parts

While the double angle trick is just that, a (useful) trick, the method of ***integration by parts*** is very general and appears in many different forms. It is the integration counterpart of the product rule for differentiation.

5.1. The product rule and integration by parts.

Recall that the product rule says that

$$\frac{\mathrm{d}F(x)G(x)}{\mathrm{d}x} = \frac{\mathrm{d}F(x)}{\mathrm{d}x}G(x) + F(x)\frac{\mathrm{d}G(x)}{\mathrm{d}x}$$

and therefore, after rearranging terms,

$$F(x)\frac{\mathrm{d}G(x)}{\mathrm{d}x} = \frac{\mathrm{d}F(x)G(x)}{\mathrm{d}x} - \frac{\mathrm{d}F(x)}{\mathrm{d}x}G(x).$$

If we integrate both sides we get the formula for ***integration by parts***

$$\int F(x)\frac{\mathrm{d}G(x)}{\mathrm{d}x}\,\mathrm{d}x = F(x)G(x) - \int \frac{\mathrm{d}F(x)}{\mathrm{d}x}G(x)\,\mathrm{d}x.$$

Note that the effect of integration by parts is to integrate one part of the function ($G'(x)$ got replaced by $G(x)$) and to differentiate the other part ($F(x)$ got replaced by $F'(x)$). For any given integral there are many ways of choosing F and G, and it not always easy to see what the best choice is.

5.2. An Example – Integrating by parts once.

Consider the problem of finding

$$I = \int xe^x\,\mathrm{d}x.$$

We can use integration by parts as follows:

$$\int \underbrace{x}_{F(x)}\underbrace{e^x}_{G'(x)}\,\mathrm{d}x = \underbrace{x}_{F(x)}\underbrace{e^x}_{G(x)} - \int \underbrace{e^x}_{G(x)}\underbrace{1}_{F'(x)}\,\mathrm{d}x = xe^x - e^x + C.$$

Observe that in this example e^x was easy to integrate, while the factor x becomes an easier function when you differentiate it. This is the usual state of affairs when integration by parts works: differentiating one of the factors ($F(x)$) should simplify the integral, while integrating the other ($G'(x)$) should not complicate things (too much).

5.3. Another example.

What is

$$\int x\sin x\,\mathrm{d}x?$$

Since $\sin x = \frac{\mathrm{d}(-\cos x)}{\mathrm{d}x}$ we can integrate by parts

$$\int \underbrace{x}_{F(x)}\underbrace{\sin x}_{G'(x)}\,\mathrm{d}x = \underbrace{x}_{F(x)}\underbrace{(-\cos x)}_{G(x)} - \int \underbrace{1}_{F'(x)}\cdot\underbrace{(-\cos x)}_{G(x)}\,\mathrm{d}x = -x\cos x + \sin x + C.$$

5.4. Example – Repeated Integration by Parts. Let's try to compute

$$I = \int x^2 e^{2x}\,\mathrm{d}x$$

by integrating by parts. Since $e^{2x} = \frac{\mathrm{d}\frac{1}{2}e^{2x}}{\mathrm{d}x}$ one has

$$\int \underbrace{x^2}_{F(x)} \underbrace{e^{2x}}_{G'(x)}\,\mathrm{d}x = x^2\frac{e^{2x}}{2} - \int \frac{e^{2x}}{2}2x\,\mathrm{d}x = \frac{1}{2}x^2e^{2x} - \int e^{2x}x\,\mathrm{d}x. \tag{3}$$

To do the integral on the left we have to integrate by parts again:

$$\int e^{2x}x\,\mathrm{d}x = \underbrace{\frac{1}{2}e^{2x}}_{G(x)}\underbrace{x}_{F(x)} - \int \underbrace{\frac{1}{2}e^{2x}}_{G(x)}\underbrace{1}_{F'(x)}\,\mathrm{d}x. = \frac{1}{2}xe^{2x} - \frac{1}{2}\int e^{2x}\,\mathrm{d}x = \frac{1}{2}xe^{2x} - \frac{1}{4}e^{2x} + C.$$

Combining this with (3) we get

$$\int x^2e^{2x}\,\mathrm{d}x = \frac{1}{2}x^2e^{2x} - \frac{1}{2}xe^{2x} + \frac{1}{4}e^{2x} - C$$

(Be careful with all the minus signs that appear when integrating by parts.)

5.5. Another example of repeated integration by parts. The same procedure as in the previous example will work whenever we have to integrate

$$\int P(x)e^{ax}\,\mathrm{d}x$$

where $P(x)$ is any polynomial, and a is a constant. Every time we integrate by parts, we get this

$$\begin{aligned}\int \underbrace{P(x)}_{F(x)}\underbrace{e^{ax}}_{G'(x)}\,\mathrm{d}x &= P(x)\frac{e^{ax}}{a} - \int \frac{e^{ax}}{a}P'(x)\,\mathrm{d}x \\ &= \frac{1}{a}P(x)e^{ax} - \frac{1}{a}\int P'(x)e^{ax}\,\mathrm{d}x.\end{aligned}$$

We have replaced the integral $\int P(x)e^{ax}\,\mathrm{d}x$ with the integral $\int P'(x)e^{ax}\,\mathrm{d}x$. This is the same kind of integral, but it is a little easier since the degree of the derivative $P'(x)$ is less than the degree of $P(x)$.

5.6. Example – sometimes the factor $G'(x)$ is invisible. Here is how we can get the antiderivative of $\ln x$ by integrating by parts:

$$\begin{aligned}\int \ln x\,\mathrm{d}x &= \int \underbrace{\ln x}_{F(x)} \cdot \underbrace{1}_{G'(x)}\,\mathrm{d}x \\ &= \ln x \cdot x - \int \frac{1}{x}\cdot x\,\mathrm{d}x \\ &= x\ln x - \int 1\,\mathrm{d}x \\ &= x\ln x - x + C.\end{aligned}$$

We can do $\int P(x)\ln x\,\mathrm{d}x$ in the same way if $P(x)$ is any polynomial. For instance, to compute

$$\int (z^2+z)\ln z\,\mathrm{d}z$$

we integrate by parts:

$$\begin{aligned}\int \underbrace{(z^2+z)}_{G'(z)}\underbrace{\ln z}_{F(z)}\,\mathrm{d}z &= \left(\tfrac{1}{3}z^3+\tfrac{1}{2}z^2\right)\ln z - \int \left(\tfrac{1}{3}z^3+\tfrac{1}{2}z^2\right)\frac{1}{z}\,\mathrm{d}z \\ &= \left(\tfrac{1}{3}z^3+\tfrac{1}{2}z^2\right)\ln z - \int \left(\tfrac{1}{3}z^2+\tfrac{1}{2}z\right)\mathrm{d}z \\ &= \left(\tfrac{1}{3}z^3+\tfrac{1}{2}z^2\right)\ln z - \tfrac{1}{9}z^3 - \tfrac{1}{4}z^2 + C.\end{aligned}$$

5.7. An example where we get the original integral back. It can happen that after integrating by parts a few times the integral we get is the same as the one we started with. When this happens we have found an equation for the integral, which we can then try to solve. The standard example in which this happens is the integral

$$I = \int e^x \sin 2x\,\mathrm{d}x.$$

We integrate by parts twice:

$$\begin{aligned}\int \underbrace{e^x}_{F'(x)}\underbrace{\sin 2x}_{G(x)}\,\mathrm{d}x &= e^x\sin 2x - \int \underbrace{e^x}_{F(x)}\underbrace{2\cos 2x}_{G'(x)}\,\mathrm{d}x \\ &= e^x\sin 2x - 2\int e^x\cos 2x\,\mathrm{d}x \\ &= e^x\sin 2x - 2e^x\cos 2x - 2\int e^x 2\sin 2x\,\mathrm{d}x \\ &= e^x\sin 2x - 2e^x\cos 2x - 4\int e^x\sin 2x\,\mathrm{d}x.\end{aligned}$$

Note that the last integral here is exactly I again. Therefore the integral I satisfies

$$I = e^x\sin 2x - 2e^x\cos 2x - 4I.$$

We solve this equation for I, with result

$$5I = e^x\sin 2x - 2e^x\cos 2x \implies I = \frac{1}{5}\left(e^x\sin 2x - 2e^x\cos 2x\right).$$

Since I is an indefinite integral we still have to add the arbitrary constant:

$$I = \frac{1}{5}\left(e^x\sin 2x - 2e^x\cos 2x\right) + C.$$

6. Reduction Formulas

We have seen that we can compute integrals by integrating by parts, and that we sometimes have to integrate by parts more than once to get the answer. There are integrals where we have to integrate by parts not once, not twice, but n-times before the answer shows up. To do such integrals it is useful to carefully describe what happens each time we integrate by parts before we do the actual integrations. The formula that describes what happens after one partial integration is called a ***reduction formula***. All this is best explained by an example.

6.1. First example of a reduction formula.

Consider the integral

$$I_n = \int x^n e^{ax}\,\mathrm{d}x, \qquad (n = 0, 1, 2, 3, \ldots)$$

or, in other words, consider all the integrals

$$I_0 = \int e^{ax}\,\mathrm{d}x, \quad I_1 = \int x e^{ax}\,\mathrm{d}x, \quad I_2 = \int x^2 e^{ax}\,\mathrm{d}x, \quad I_3 = \int x^3 e^{ax}\,\mathrm{d}x, \ldots$$

and so on. We will consider all these integrals at the same time.

Integration by parts in I_n gives us

$$\begin{aligned} I_n &= \int \underbrace{x^n}_{F(x)} \underbrace{e^{ax}}_{G'(x)}\,\mathrm{d}x \\ &= x^n \frac{1}{a} e^{ax} - \int n x^{n-1} \frac{1}{a} e^{ax}\,\mathrm{d}x \\ &= \frac{1}{a} x^n e^{ax} - \frac{n}{a} \int x^{n-1} e^{ax}\,\mathrm{d}x. \end{aligned}$$

We haven't computed the integral, and in fact the integral that we still have to do is of the same kind as the one we started with (integral of $x^{n-1}e^{ax}$ instead of $x^n e^{ax}$). What we have derived is the following ***reduction formula***

$$I_n = \frac{1}{a} x^n e^{ax} - \frac{n}{a} I_{n-1},$$

which holds for all n.

For $n = 0$ we do not need the reduction formula to find the integral. We have

$$I_0 = \int e^{ax}\,\mathrm{d}x = \frac{1}{a} e^{ax} + C.$$

When $n \neq 0$ the reduction formula tells us that we have to compute I_{n-1} if we want to find I_n. The point of a reduction formula is that the same formula also applies to I_{n-1}, and I_{n-2}, etc., so that after repeated application of the formula we end up with I_0, i.e., an integral we know.

For example, if we want to compute $\int x^3 e^{ax}\,\mathrm{d}x$ we use the reduction formula three times:

$$\begin{aligned} I_3 &= \frac{1}{a} x^3 e^{ax} - \frac{3}{a} I_2 \\ &= \frac{1}{a} x^3 e^{ax} - \frac{3}{a} \left\{ \frac{1}{a} x^2 e^{ax} - \frac{2}{a} I_1 \right\} \\ &= \frac{1}{a} x^3 e^{ax} - \frac{3}{a} \left\{ \frac{1}{a} x^2 e^{ax} - \frac{2}{a} \left(\frac{1}{a} x e^{ax} - \frac{1}{a} I_0 \right) \right\} \end{aligned}$$

Insert the known integral $I_0 = \frac{1}{a} e^{ax} + C$ and simplify the other terms and we get

$$\int x^3 e^{ax}\,\mathrm{d}x = \frac{1}{a} x^3 e^{ax} - \frac{3}{a^2} x^2 e^{ax} + \frac{6}{a^3} x e^{ax} - \frac{6}{a^4} e^{ax} + C.$$

6.2. Reduction formula requiring two partial integrations. Consider

$$S_n = \int x^n \sin x \, \mathrm{d}x.$$

Then for $n \geq 2$ one has

$$\begin{aligned} S_n &= -x^n \cos x + n \int x^{n-1} \cos x \, \mathrm{d}x \\ &= -x^n \cos x + n x^{n-1} \sin x - n(n-1) \int x^{n-2} \sin x \, \mathrm{d}x. \end{aligned}$$

Thus we find the reduction formula

$$S_n = -x^n \cos x + n x^{n-1} \sin x - n(n-1) S_{n-2}.$$

Each time we use this reduction, the exponent n drops by 2, so in the end we get either S_1 or S_0, depending on whether we started with an odd or even n. These two integrals are

$$\begin{aligned} S_0 &= \int \sin x \, \mathrm{d}x = -\cos x + C \\ S_1 &= \int x \sin x \, \mathrm{d}x = -x \cos x + \sin x + C. \end{aligned}$$

(Integrate by parts once to find S_1.)

As an example of how to use the reduction formulas for S_n let's try to compute S_4:

$$\begin{aligned} \int x^4 \sin x \, \mathrm{d}x = S_4 &= -x^4 \cos x + 4x^3 \sin x - 4 \cdot 3 S_2 \\ &= -x^4 \cos x + 4x^3 \sin x - 4 \cdot 3 \cdot \left\{ -x^2 \cos x + 2x \sin x - 2 \cdot 1 S_0 \right\} \end{aligned}$$

At this point we use $S_0 = \int \sin x \, \mathrm{d}x = -\cos x + C$, and we combine like terms. This results in

$$\begin{aligned} \int x^4 \sin x \, \mathrm{d}x &= -x^4 \cos x + 4x^3 \sin x \\ &\qquad - 4 \cdot 3 \cdot \left\{ -x^2 \cos x + 2x \sin x - 2 \cdot 1(-\cos x) \right\} + C \\ &= \left(-x^4 + 12x^2 - 24 \right) \cos x + \left(4x^3 + 24x \right) \sin x + C. \end{aligned}$$

6.3. A reduction formula where you have to solve for I_n. We try to compute

$$I_n = \int (\sin x)^n \, \mathrm{d}x$$

by a reduction formula. Integrating by parts we get

$$\begin{aligned} I_n &= \int (\sin x)^{n-1} \sin x \, \mathrm{d}x \\ &= -(\sin x)^{n-1} \cos x - \int (-\cos x)(n-1)(\sin x)^{n-2} \cos x \, \mathrm{d}x \\ &= -(\sin x)^{n-1} \cos x + (n-1) \int (\sin x)^{n-2} \cos^2 x \, \mathrm{d}x. \end{aligned}$$

We now use $\cos^2 x = 1 - \sin^2 x$, which gives

$$\begin{aligned} I_n &= -(\sin x)^{n-1}\cos x + (n-1)\int \left\{\sin^{n-2} x - \sin^n x\right\}\,\mathrm{d}x \\ &= -(\sin x)^{n-1}\cos x + (n-1)I_{n-2} - (n-1)I_n. \end{aligned}$$

We can think of this as an equation for I_n, which, when we solve it tells us

$$nI_n = -(\sin x)^{n-1}\cos x + (n-1)I_{n-2}$$

and thus implies

$$I_n = -\frac{1}{n}\sin^{n-1} x\cos x + \frac{n-1}{n}I_{n-2}. \tag{4}$$

Since we know the integrals

$$I_0 = \int (\sin x)^0 \mathrm{d}x = \int \mathrm{d}x = x + C$$

and

$$I_1 = \int \sin x\,\mathrm{d}x = -\cos x + C$$

the reduction formula (4) allows us to calculate I_n for any $n \ge 2$.

6.4. A reduction formula that will come in handy later. In the next section we will see how the integral of any "rational function" can be transformed into integrals of easier functions, the most difficult of which turns out to be

$$I_n = \int \frac{\mathrm{d}x}{(1+x^2)^n}.$$

When $n = 1$ this is a standard integral, namely

$$I_1 = \int \frac{\mathrm{d}x}{1+x^2} = \arctan x + C.$$

When $n > 1$ integration by parts gives us a reduction formula. Here's the computation:

$$\begin{aligned} I_n &= \int (1+x^2)^{-n}\,\mathrm{d}x \\ &= \frac{x}{(1+x^2)^n} - \int x\,(-n)\big(1+x^2\big)^{-n-1}\,2x\,\mathrm{d}x \\ &= \frac{x}{(1+x^2)^n} + 2n\int \frac{x^2}{(1+x^2)^{n+1}}\,\mathrm{d}x \end{aligned}$$

Apply

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

to get

$$\int \frac{x^2}{(1+x^2)^{n+1}}\,\mathrm{d}x = \int \left\{\frac{1}{(1+x^2)^n} - \frac{1}{(1+x^2)^{n+1}}\right\}\mathrm{d}x = I_n - I_{n+1}.$$

Our integration by parts therefore told us that

$$I_n = \frac{x}{(1+x^2)^n} + 2n\big(I_n - I_{n+1}\big),$$

which we can solve for I_{n+1}. We find the reduction formula

$$I_{n+1} = \frac{1}{2n}\frac{x}{(1+x^2)^n} + \frac{2n-1}{2n} I_n.$$

As an example of how we can use it, we start with $I_1 = \arctan x + C$, and conclude that

$$\begin{aligned}
\int \frac{\mathrm{d}x}{(1+x^2)^2} &= I_2 = I_{1+1} \\
&= \frac{1}{2\cdot 1}\frac{x}{(1+x^2)^1} + \frac{2\cdot 1 - 1}{2\cdot 1} I_1 \\
&= \tfrac{1}{2}\frac{x}{1+x^2} + \tfrac{1}{2}\arctan x + C.
\end{aligned}$$

Apply the reduction formula again, now with $n = 2$, and we get

$$\begin{aligned}
\int \frac{\mathrm{d}x}{(1+x^2)^3} &= I_3 = I_{2+1} \\
&= \frac{1}{2\cdot 2}\frac{x}{(1+x^2)^2} + \frac{2\cdot 2 - 1}{2\cdot 2} I_2 \\
&= \tfrac{1}{4}\frac{x}{(1+x^2)^2} + \tfrac{3}{4}\left\{\tfrac{1}{2}\frac{x}{1+x^2} + \tfrac{1}{2}\arctan x\right\} \\
&= \tfrac{1}{4}\frac{x}{(1+x^2)^2} + \tfrac{3}{8}\frac{x}{1+x^2} + \tfrac{3}{8}\arctan x + C.
\end{aligned}$$

7. Problems

1. Evaluate $\displaystyle\int x^n \ln x \,\mathrm{d}x$ where $n \neq -1$.

2. Assume a and b are constants, and compute $\displaystyle\int e^{ax}\sin bx\,\mathrm{d}x$. [Hint: Integrate by parts twice; you can assume that $b \neq 0$.]

3. Evaluate $\displaystyle\int e^{ax}\cos bx\,\mathrm{d}x$ where $a, b \neq 0$.

4. Prove the formula

$$\int x^n e^x\,\mathrm{d}x = x^n e^x - n\int x^{n-1}e^x\,\mathrm{d}x$$

and use it to evaluate $\displaystyle\int x^2 e^x\,\mathrm{d}x$.

5. Use §6.3 to evaluate $\displaystyle\int \sin^2 x\,\mathrm{d}x$. Show that the answer is the same as the answer you get using the half angle formula.

6. Use the reduction formula in §6.3 to compute $\displaystyle\int_0^{\pi/2}\sin^{14} x\mathrm{d}x$.

7. In this problem you'll look at the numbers

$$A_n = \int_0^{\pi}\sin^n x\,\mathrm{d}x.$$

(a) Check that $A_0 = \pi$ and $A_1 = 2$.

(b) Use the reduction formula in §6.3 to compute A_5, A_6, and A_7.

(c) Explain why

$$A_n < A_{n-1}$$

is true for all $n = 1, 2, 3, 4, \ldots$

(Hint: Interpret the integrals A_n as area under the graph, and check that $(\sin x)^n \leq (\sin x)^{n-1}$ for all x.)

(d) Based on your values for A_5, A_6, and A_7 find two fractions a and b such that $a < \pi < b$.

8. Prove the formula

$$\int \cos^n x\,\mathrm{d}x = \frac{1}{n}\sin x\cos^{n-1}x + \frac{n-1}{n}\int\cos^{n-2}x\,\mathrm{d}x,$$

for $n \neq 0$, and use it to evaluate

$$\int_0^{\pi/4}\cos^4 x\,\mathrm{d}x.$$

9. Prove the formula

$$\int x^m (\ln x)^n \, dx = \frac{x^{m+1}(\ln x)^n}{m+1} - \frac{n}{m+1}\int x^m (\ln x)^{n-1}\, dx,$$

for $m \neq -1$, and use it to evaluate the following integrals:

10. $\int \ln x \, dx$

11. $\int (\ln x)^2 \, dx$

12. $\int x^3 (\ln x)^2 \, dx$

13. Evaluate $\int x^{-1} \ln x \, dx$ by another method. [Hint: the solution is short!]

14. For any integer $n > 1$ derive the formula

$$\int \tan^n x \, dx = \frac{\tan^{n-1} x}{n-1} - \int \tan^{n-2} x \, dx$$

Using this, find $\int_0^{\pi/4} \tan^5 x \, dx$.

Use the reduction formula from example 6.4 to compute these integrals:

15. $\int \frac{dx}{(1+x^2)^3}$

16. $\int \frac{dx}{(1+x^2)^4}$

17. $\int \frac{x\,dx}{(1+x^2)^4}$ [Hint: $\int x/(1+x^2)^n dx$ is easy.]

18. $\int \frac{1+x}{(1+x^2)^2}\, dx$

19. $\int \frac{dx}{\left(49+x^2\right)^3}$.

20. The reduction formula from example 6.4 is valid for all $n \neq 0$. In particular, n does not have to be an integer, and it does not have to be positive. Find a relation between $\int \sqrt{1+x^2}\, dx$ and $\int \frac{dx}{\sqrt{1+x^2}}$ by setting $n = -\frac{1}{2}$.

21. Apply integration by parts to

$$\int \frac{1}{x}\, dx$$

Let $u = \frac{1}{x}$ and $dv = dx$. This gives us, $du = \frac{-1}{x^2}\, dx$ and $v = x$.

$$\int \frac{1}{x}\, dx = (\frac{1}{x})(x) - \int x \frac{-1}{x^2}\, dx$$

Simplifying

$$\int \frac{1}{x}\, dx = 1 + \int \frac{1}{x}\, dx$$

and subtracting the integral from both sides gives us $0 = 1$. How can this be?

8. Partial Fraction Expansion

By definition, a ***rational function*** is a quotient (a *ratio*) of polynomials,

$$f(x) = \frac{P(x)}{Q(x)} = \frac{p_n x^n + p_{n-1}x^{n-1} + \cdots + p_1 x + p_0}{q_d x^d + q_{d-1}x^{d-1} + \cdots + q_1 x + q_0}.$$

Such rational functions can always be integrated, and the trick that allows you to do this is called a ***partial fraction expansion***. The whole procedure consists of several steps that are explained in this section. The procedure itself has nothing to do with integration: it's just a way of rewriting rational functions. It is in fact useful in other situations, such as finding Taylor expansions (see Chapter 4) and computing "inverse Laplace transforms" (see MATH 319.)

8.1. Reduce to a proper rational function.

A ***proper rational function*** is a rational function $P(x)/Q(x)$ where the degree of $P(x)$ is strictly less than the degree of $Q(x)$. The method of partial fractions only applies to proper rational functions. Fortunately there's an additional trick for dealing with rational functions that are not proper.

If P/Q isn't proper, i.e. if degree$(P) \geq$ degree(Q), then you divide P by Q, with result

$$\frac{P(x)}{Q(x)} = S(x) + \frac{R(x)}{Q(x)}$$

where $S(x)$ is the quotient, and $R(x)$ is the remainder after division. In practice you would do a long division to find $S(x)$ and $R(x)$.

8.2. Example. Consider the rational function

$$f(x) = \frac{x^3 - 2x + 2}{x^2 - 1}.$$

Here the numerator has degree 3 which is more than the degree of the denominator (which is 2). The function $f(x)$ is therefore not a proper rational function. To apply the method of partial fractions we must first do a division with remainder. One has

$$\begin{array}{r|l} & \quad x \qquad\quad = S(x) \\ x^2 - 1 & x^3 - 2x + 2 \\ & \underline{x^3 \;\; -x} \\ & \qquad -x + 2 = R(x) \end{array}$$

so that

$$f(x) = \frac{x^3 - 2x + 2}{x^2 - 1} = x + \frac{-x + 2}{x^2 - 1}$$

When we integrate we get

$$\begin{aligned} \int \frac{x^3 - 2x + 2}{x^2 - 1}\,\mathrm{d}x &= \int \left\{ x + \frac{-x + 2}{x^2 - 1} \right\} \mathrm{d}x \\ &= \frac{x^2}{2} + \int \frac{-x + 2}{x^2 - 1}\,\mathrm{d}x. \end{aligned}$$

The rational function that we still have to integrate, namely $\frac{-x+2}{x^2-1}$, is proper: its numerator has lower degree than its denominator.

8.3. Partial Fraction Expansion: The Easy Case. To compute the partial fraction expansion of a proper rational function $P(x)/Q(x)$ you must factor the denominator $Q(x)$. Factoring the denominator is a problem as difficult as finding all of its roots; in Math 222 we shall only do problems where the denominator is already factored into linear and quadratic factors, or where this factorization is easy to find.

In the easiest partial fractions problems, all the roots of $Q(x)$ are real numbers and distinct, so the denominator is factored into distinct linear factors, say

$$\frac{P(x)}{Q(x)} = \frac{P(x)}{(x - a_1)(x - a_2)\cdots(x - a_n)}.$$

To integrate this function we find constants $A_1, A_2, \ldots, A_n$ so that

$$\frac{P(x)}{Q(x)} = \frac{A_1}{x - a_1} + \frac{A_2}{x - a_2} + \cdots + \frac{A_n}{x - a_n}. \tag{\#}$$

Then the integral is

$$\int \frac{P(x)}{Q(x)}\,\mathrm{d}x = A_1 \ln|x - a_1| + A_2 \ln|x - a_2| + \cdots + A_n \ln|x - a_n| + C.$$

One way to find the coefficients A_i in (#) is called the ***method of equating coefficients***. In this method we multiply both sides of (#) with $Q(x) = (x - a_1)\cdots(x - a_n)$.

The result is a polynomial of degree n on both sides. Equating the coefficients of these polynomial gives a system of n linear equations for A_1, ..., A_n. You get the A_i by solving that system of equations.

Another much faster way to find the coefficients A_i is the ***Heaviside trick***[1]. Multiply equation (#) by $x - a_i$ and then plug in[2] $x = a_i$. On the right you are left with A_i so

$$A_i = \frac{P(x)(x-a_i)}{Q(x)}\bigg|_{x=a_i} = \frac{P(a_i)}{(a_i - a_1)\cdots(a_i - a_{i-1})(a_i - a_{i+1})\cdots(a_i - a_n)}.$$

8.4. Previous Example continued. To integrate $\dfrac{-x+2}{x^2-1}$ we factor the denominator,

$$x^2 - 1 = (x-1)(x+1).$$

The partial fraction expansion of $\dfrac{-x+2}{x^2-1}$ then is

$$\frac{-x+2}{x^2-1} = \frac{-x+2}{(x-1)(x+1)} = \frac{A}{x-1} + \frac{B}{x+1}. \tag{5}$$

Multiply with $(x-1)(x+1)$ to get

$$-x + 2 = A(x+1) + B(x-1) = (A+B)x + (A-B).$$

The functions of x on the left and right are equal only if the coefficient of x and the constant term are equal. In other words we must have

$$A + B = -1 \quad \text{and} \quad A - B = 2.$$

These are two linear equations for two unknowns A and B, which we now proceed to solve. Adding both equations gives $2A = 1$, so that $A = \frac{1}{2}$; from the first equation one then finds $B = -1 - A = -\frac{3}{2}$. So

$$\frac{-x+2}{x^2-1} = \frac{1/2}{x-1} - \frac{3/2}{x+1}.$$

Instead, we could also use the Heaviside trick: multiply (5) with $x - 1$ to get

$$\frac{-x+2}{x+1} = A + B\frac{x-1}{x+1}$$

Take the limit $x \to 1$ and you find

$$\frac{-1+2}{1+1} = A, \text{ i.e. } A = \frac{1}{2}.$$

Similarly, after multiplying (5) with $x + 1$ one gets

$$\frac{-x+2}{x-1} = A\frac{x+1}{x-1} + B,$$

and letting $x \to -1$ you find

$$B = \frac{-(-1)+2}{(-1)-1} = -\frac{3}{2},$$

as before.

[1] Named after Oliver Heaviside, a physicist and electrical engineer in the late 19th and early 20th century.

[2] More properly, you should take the limit $x \to a_i$. The problem here is that equation (#) has $x - a_i$ in the denominator, so that it does not hold for $x = a_i$. Therefore you cannot set x equal to a_i in any equation derived from (#). But you can take the limit $x \to a_i$, which in practice is just as good.

Either way, the integral is now easily found, namely,

$$\begin{aligned}\int \frac{x^3-2x+1}{x^2-1}\,\mathrm{d}x &= \frac{x^2}{2} + \int \frac{-x+2}{x^2-1}\,\mathrm{d}x \\ &= \frac{x^2}{2} + \int \left\{\frac{1/2}{x-1} - \frac{3/2}{x+1}\right\}\mathrm{d}x \\ &= \frac{x^2}{2} + \frac{1}{2}\ln|x-1| - \frac{3}{2}\ln|x+1| + C.\end{aligned}$$

8.5. Partial Fraction Expansion: The General Case. When the denominator $Q(x)$ contains repeated factors or quadratic factors (or both) the partial fraction decomposition is more complicated. In the most general case the denominator $Q(x)$ can be factored in the form

$$Q(x) = (x-a_1)^{k_1}\cdots(x-a_n)^{k_n}(x^2+b_1x+c_1)^{\ell_1}\cdots(x^2+b_mx+c_m)^{\ell_m} \tag{6}$$

Here we assume that the factors $x-a_1$, ..., $x-a_n$ are all different, and we also assume that the factors $x^2+b_1x+c_1$, ..., $x^2+b_mx+c_m$ are all different.

It is a theorem from advanced algebra that you can always write the rational function $P(x)/Q(x)$ as a sum of terms like this

$$\frac{P(x)}{Q(x)} = \cdots + \frac{A}{(x-a_i)^k} + \cdots + \frac{Bx+C}{(x^2+b_jx+c_j)^\ell} + \cdots \tag{7}$$

How did this sum come about?

For each linear factor $(x-a)^k$ in the denominator (6) you get terms

$$\frac{A_1}{x-a} + \frac{A_2}{(x-a)^2} + \cdots + \frac{A_k}{(x-a)^k}$$

in the decomposition. There are as many terms as the exponent of the linear factor that generated them.

For each quadratic factor $(x^2+bx+c)^\ell$ you get terms

$$\frac{B_1x+C_1}{x^2+bx+c} + \frac{B_2x+C_2}{(x^2+bx+c)^2} + \cdots + \frac{B_mx+C_m}{(x^2+bx+c)^\ell}.$$

Again, there are as many terms as the exponent ℓ with which the quadratic factor appears in the denominator (6).

In general, you find the constants $A_{\ldots}$, $B_{\ldots}$ and $C_{\ldots}$ by the method of equating coefficients.

Unfortunately, in the presence of quadratic factors or repeated linear factors the Heaviside trick does not give the whole answer; we really have to use the method of equating coefficients.

The workings of this method are best explained in an example.

8.6. Example. Find the partial fraction decomposition of

$$f(x) = \frac{x^2+2}{x^2(x^2+1)}$$

and compute

$$I = \int \frac{x^2+2}{x^2(x^2+1)}\,\mathrm{d}x.$$

The degree of the denominator $x^2(x^2+1)$ is four, so our partial fraction decomposition must also contain four undetermined constants. The expansion should be of the form

$$\frac{x^2+2}{x^2(x^2+1)} = \frac{A}{x} + \frac{B}{x^2} + \frac{Cx+D}{x^2+1}.$$

To find the coefficients A, B, C, D we multiply both sides with $x^2(1+x^2)$,

$$\begin{aligned} x^2+2 &= Ax(x^2+1) + B(x^2+1) + x^2(Cx+D) \\ x^2+2 &= (A+C)x^3 + (B+D)x^2 + Ax + B \\ 0\cdot x^3 + 1\cdot x^2 + 0\cdot x + 2 &= (A+C)x^3 + (B+D)x^2 + Ax + B \end{aligned}$$

Comparing terms with the same power of x we find that

$$A+C=0, \quad B+D=1, \quad A=0, \quad B=2.$$

These are four equations for four unknowns. Fortunately for us they are not very difficult in this example. We find $A=0$, $B=2$, $C=-A=0$, and $D=1-B=-1$, whence

$$f(x) = \frac{x^2+2}{x^2(x^2+1)} = \frac{2}{x^2} - \frac{1}{x^2+1}.$$

The integral is therefore

$$I = \frac{x^2+2}{x^2(x^2+1)}\,\mathrm{d}x = -\frac{2}{x} - \arctan x + C.$$

8.7. A complicated example. Find the integral

$$\int \frac{x^2+3}{x^2(x+1)(x^2+1)^3}\,\mathrm{d}x.$$

The procedure is exactly the same as in the previous example. We have to expand the integrand in partial fractions:

$$\begin{aligned} \frac{x^2+3}{x^2(x+1)(x^2+1)^3} &= \frac{A_1}{x} + \frac{A_2}{x^2} + \frac{A_3}{x+1} \\ &\quad + \frac{B_1x+C_1}{x^2+1} + \frac{B_2x+C_2}{(x^2+1)^2} + \frac{B_3x+C_3}{(x^2+1)^3}. \end{aligned} \tag{8}$$

Note that the degree of the denominator $x^2(x+1)(x^2+1)^3$ is $2+1+3\times 2=9$, and also that the partial fraction decomposition has nine undetermined constants A_1, A_2, A_3, B_1, C_1, B_2, C_2, B_3, C_3. After multiplying both sides of (8) with the denominator $x^2(x+1)(x^2+1)^3$, expanding everything, and then equating coefficients of powers of x on both sides, we get a system of nine linear equations in these nine unknowns. The final step in finding the partial fraction decomposition is to solve those linear equations. A computer program like `Maple` or `Mathematica` can do this easily, but it is a lot of work to do it by hand.

8.8. After the partial fraction decomposition. Once we have the partial fraction decomposition (8) we still have to integrate the terms that appeared. The first three terms are of the form $\int A(x-a)^{-p}\,\mathrm{d}x$ and they are easy to integrate:

$$\int \frac{A\,\mathrm{d}x}{x-a} = A\ln|x-a| + C$$

and

$$\int \frac{A\,\mathrm{d}x}{(x-a)^p} = \frac{A}{(1-p)(x-a)^{p-1}} + C$$

if $p > 1$. The next, fourth term in (8) can be written as

$$\begin{aligned}\int \frac{B_1 x + C_1}{x^2+1}\,\mathrm{d}x &= B_1 \int \frac{x}{x^2+1}\,\mathrm{d}x + C_1 \int \frac{\mathrm{d}x}{x^2+1} \\ &= \frac{B_1}{2}\ln(x^2+1) + C_1 \arctan x + K,\end{aligned}$$

where K is the integration constant (normally "$+C$" but there are so many other C's in this problem that we chose a different letter, just for this once.)

While these integrals are already not very simple, the integrals

$$\int \frac{Bx + C}{(x^2+bx+c)^p}\,\mathrm{d}x \qquad \text{with } p > 1$$

which can appear are particularly unpleasant. If we really must compute one of these, then we should first complete the square in the denominator so that the integral takes the form

$$\int \frac{Ax + B}{((x+b)^2 + a^2)^p}\,\mathrm{d}x.$$

After the change of variables $u = x + b$ and factoring out constants we are left with the integrals

$$\int \frac{du}{(u^2+a^2)^p} \qquad \text{and} \qquad \int \frac{u\,\mathrm{d}u}{(u^2+a^2)^p}.$$

The reduction formula from example 6.4 then allows us to compute this integral.

An alternative approach is to use complex numbers. If we allow complex numbers then the quadratic factors $x^2 + bx + c$ can be factored, and our partial fraction expansion only contains terms of the form $A/(x-a)^p$, although A and a can now be complex numbers. The integrals are then easy, but the answer has complex numbers in it, and rewriting the answer in terms of real numbers again can be quite involved. In this course we will avoid complex numbers and therefore we will not explain this any further.

9. Problems

1. Express each of the following rational functions as a polynomial plus a proper rational function. (See §8.1 for definitions.)

(a) $\dfrac{x^3}{x^3-4}$

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

(c) $\dfrac{x^3-x^2-x-5}{x^3-4}$

(d) $\dfrac{x^3-1}{x^2-1}$

2. Compute the following integrals by completing the square:

(a) $\displaystyle\int \frac{\mathrm{d}x}{x^2+6x+8}$,

(b) $\displaystyle\int \frac{\mathrm{d}x}{x^2+6x+10}$,

(c) $\displaystyle\int \frac{\mathrm{d}x}{5x^2+20x+25}$.

3. Use the method of equating coefficients to find numbers A, B, C such that

$$\frac{x^2+3}{x(x+1)(x-1)} = \frac{A}{x} + \frac{B}{x+1} + \frac{C}{x-1}$$

and then evaluate the integral

$$\int \frac{x^2+3}{x(x+1)(x-1)}\,\mathrm{d}x.$$

4. Do the previous problem using the Heaviside trick.

5. Find the integral $\displaystyle\int \frac{x^2+3}{x^2(x-1)}\,\mathrm{d}x$.

6. Simplicio had to integrate
$$\frac{4x^2}{(x-3)(x+1)}.$$
He set
$$\frac{4x^2}{(x-3)(x+1)} = \frac{A}{x-3} + \frac{B}{x+1}.$$
Using the Heaviside trick he then found
$$A = \left.\frac{4x^2}{x-3}\right|_{x=-1} = -1,$$
and
$$B = \left.\frac{4x^2}{x+1}\right|_{x=3} = 9,$$
which leads him to conclude that
$$\frac{4x^2}{(x-3)(x+1)} = \frac{-1}{x-3} + \frac{9}{x+1}.$$
To double check he now sets $x = 0$ which leads to
$$0 = \frac{1}{3} + 9 \quad ????$$
What went wrong?

Evaluate the following integrals:

7. $\displaystyle\int_{-5}^{-2} \frac{x^4-1}{x^2+1}\,\mathrm{d}x$

8. $\displaystyle\int \frac{x^3\,\mathrm{d}x}{x^4+1}$

9. $\displaystyle\int \frac{x^5\,\mathrm{d}x}{x^2-1}$

10. $\displaystyle\int \frac{x^5\,\mathrm{d}x}{x^4-1}$

11. $\displaystyle\int \frac{x^3}{x^2-1}\,\mathrm{d}x$

12. $\displaystyle\int \frac{2x+1}{x^2-3x+2}\,\mathrm{d}x$

13. $\displaystyle\int \frac{x^2+1}{x^2-3x+2}\,\mathrm{d}x$

14. $\displaystyle\int \frac{e^{3x}\,\mathrm{d}x}{e^{4x}-1}$

15. $\displaystyle\int \frac{e^x\,\mathrm{d}x}{\sqrt{1+e^{2x}}}$

16. $\displaystyle\int \frac{e^x\,\mathrm{d}x}{e^{2x}+2e^x+2}$

17. $\displaystyle\int \frac{\mathrm{d}x}{1+e^x}$

18. $\displaystyle\int \frac{\mathrm{d}x}{x(x^2+1)}$

19. $\displaystyle\int \frac{\mathrm{d}x}{x(x^2+1)^2}$

20. $\displaystyle\int \frac{\mathrm{d}x}{x^2(x-1)}$

21. $\displaystyle\int \frac{1}{(x-1)(x-2)(x-3)}\,\mathrm{d}x$

22. $\displaystyle\int \frac{x^2+1}{(x-1)(x-2)(x-3)}\,\mathrm{d}x$

23. $\displaystyle\int \frac{x^3+1}{(x-1)(x-2)(x-3)}\,\mathrm{d}x$

24. **(a)** Compute $\displaystyle\int_1^2 \frac{\mathrm{d}x}{x(x-h)}$ where h is a positive number.

(b) What happens to your answer to (a) when $h \searrow 0$?

(c) Compute $\displaystyle\int_1^2 \frac{\mathrm{d}x}{x^2}$.

10. Substitutions for integrals containing the expression $\sqrt{ax^2+bx+c}$

The main method for finding antiderivatives that we saw in Math 221 is the *method of substitution*. This method will only let us compute an integral if we happen to guess the right substitution, and guessing the right substitution is often not easy. If the integral contains the square root of a linear or quadratic function, then there are a number of substitutions that are known to help.

- Integrals with $\sqrt{ax+b}$: substitute $ax+b=u^2$ with $u>0$. See § 10.1.
- Integrals with $\sqrt{ax^2+bx+c}$: first complete the square to reduce the integral to one containing one of the following three forms
$$\sqrt{1-u^2}, \quad \sqrt{u^2-1}, \quad \sqrt{u^2+1}.$$

Then, depending on which of these three cases presents itself, you choose an appropriate substitution. There are several options:

- a trigonometric substitution; this works well in some cases, but often you end up with an integral containing trigonometric functions that is still not easy (see § 10.2 and § 10.4.1).
- use hyperbolic functions; the *hyperbolic sine* and *hyperbolic cosine* sometimes let you handle cases where trig substitutions do not help.
- a rational substitution (see § 11) using the two functions $U(t) = \frac{1}{2}\big(t + t^{-1}\big)$ and $V(t) = \frac{1}{2}\big(t - t^{-1}\big)$.

10.1. Integrals involving $\sqrt{ax+b}$. If an integral contains the square root of a linear function, i.e. $\sqrt{ax+b}$ then you can remove this square root by substituting $u = \sqrt{ax+b}$.

10.1.1. *Example.* To compute

$$I = \int x\sqrt{2x+3}\,\mathrm{d}x$$

we substitute $u = \sqrt{2x+3}$. Then

$$x = \frac{1}{2}(u^2 - 3) \text{ so that } \mathrm{d}x = u\,\mathrm{d}u,$$

and hence

$$\begin{aligned} I &= \int \underbrace{\tfrac{1}{2}(u^2-3)}_{x}\ \underbrace{u}_{\sqrt{2x+3}}\ \underbrace{u\,\mathrm{d}u}_{\mathrm{d}x} \\ &= \frac{1}{2}\int \big(u^4 - 3u^2\big)\,\mathrm{d}u \\ &= \frac{1}{2}\big\{\tfrac{1}{5}u^5 - u^3\big\} + C. \end{aligned}$$

To write the antiderivative in terms of the original variable you substitute $u = \sqrt{2x+3}$ again, which leads to

$$I = \frac{1}{10}(2x+3)^{5/2} - \frac{1}{2}(2x+3)^{3/2} + C.$$

A comment: setting $u = \sqrt{ax+b}$ is usually the best choice, but sometimes other choices also work. You, the reader, might want to try this same example substituting $v = 2x+3$ instead of the substitution we used above. You should of course get the same answer.

10.1.2. *Another example.* Compute

$$I = \int \frac{\mathrm{d}x}{1+\sqrt{1+x}}.$$

Again we substitute $u^2 = x+1$, or, $u = \sqrt{x+1}$. We get

$$\begin{aligned} I &= \int \frac{\mathrm{d}x}{1+\sqrt{1+x}} && u^2 = x+1 \text{ so } 2u\,\mathrm{d}u = \mathrm{d}x \\ &= \int \frac{2u\,\mathrm{d}u}{1+u} && \text{A rational function: we know what to do.} \\ &= \int \Big(2 - \frac{2}{1+u}\Big)\,\mathrm{d}u \\ &= 2u - 2\ln(1+u) + C \\ &= 2\sqrt{x+1} - 2\ln\big(1+\sqrt{x+1}\big) + C. \end{aligned}$$

Note that $u = \sqrt{x+1}$ is positive, so that $1+\sqrt{x+1} > 0$, and so that we do not need absolute value signs in $\ln(1+u)$.

10.2. Integrals containing $\sqrt{1-x^2}$. If an integral contains the expression $\sqrt{1-x^2}$ then this expression can be removed at the expense of introducing trigonometric functions. Sometimes (but not always) the resulting integral is easier.

The substitution that removes the $\sqrt{1-x^2}$ is $x = \sin\theta$.

10.2.1. *Example.* To compute

$$I = \int \frac{\mathrm{d}x}{(1-x^2)^{3/2}}$$

note that

$$\frac{1}{(1-x^2)^{3/2}} = \frac{1}{(1-x^2)\sqrt{1-x^2}},$$

so we have an integral involving $\sqrt{1-x^2}$.

We set $x = \sin\theta$, and thus $\mathrm{d}x = \cos\theta\,\mathrm{d}\theta$. We get

$$I = \int \frac{\cos\theta\,\mathrm{d}\theta}{(1-\sin^2\theta)^{3/2}}.$$

Use $1-\sin^2\theta = \cos^2\theta$ and you get

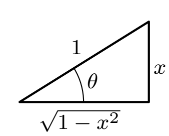
$$(1-\sin^2\theta)^{3/2} = \big(\cos^2\theta\big)^{3/2} = |\cos\theta|^3.$$

We were forced to include the absolute values here because of the possibility that $\cos\theta$ might be negative. However it turns out that $\cos\theta > 0$ in our situation since, in the original integral I the variable x must lie between -1 and $+1$: hence, if we set $x = \sin\theta$, then we may assume that $-\frac{\pi}{2} < \theta < \frac{\pi}{2}$. For those θ one has $\cos\theta > 0$, and therefore we can write

$$(1-\sin^2\theta)^{3/2} = \cos^3\theta.$$

After substitution our integral thus becomes

$$I = \int \frac{\cos\theta\,\mathrm{d}\theta}{\cos^3\theta} = \int \frac{\mathrm{d}\theta}{\cos^2\theta} = \tan\theta + C.$$



$x = \sin\theta$
$\sqrt{1-x^2} = \cos\theta.$

To express the antiderivative in terms of the original variable we use

$$x = \sin\theta \implies \tan\theta = \frac{x}{\sqrt{1-x^2}}.$$

The final result is

$$I = \int \frac{\mathrm{d}x}{(1-x^2)^{3/2}} = \frac{x}{\sqrt{1-x^2}} + C.$$

10.2.2. *Example: sometimes you don't have to do a trig substitution.* The following integral is very similar to the one from the previous example:

$$\tilde{I} = \int \frac{x\,\mathrm{d}x}{\left(1-x^2\right)^{3/2}}.$$

The only difference is an extra "x" in the numerator.

To compute this integral you can substitute $u = 1 - x^2$, in which case $\mathrm{d}u = -2x\,\mathrm{d}x$. Thus we find

$$\begin{aligned}\int \frac{x\,\mathrm{d}x}{\left(1-x^2\right)^{3/2}} &= -\frac{1}{2}\int \frac{\mathrm{d}u}{u^{3/2}} = -\frac{1}{2}\int u^{-3/2}\mathrm{d}u \\ &= -\frac{1}{2}\,\frac{u^{-1/2}}{(-1/2)} + C = \frac{1}{\sqrt{u}} + C \\ &= \frac{1}{\sqrt{1-x^2}} + C.\end{aligned}$$

10.3. Integrals containing $\sqrt{a^2-x^2}$. If an integral contains the expression $\sqrt{a^2-x^2}$ for some positive number a, then this can be removed by substituting either $x = a\sin\theta$ or $x = a\cos\theta$. Since in the integral we must have $-a < x < a$, we only need values of θ in the interval $(-\frac{\pi}{2}, \frac{\pi}{2})$. Thus we substitute

$$x = a\sin\theta, \quad -\frac{\pi}{2} < \theta < \frac{\pi}{2}.$$

For these values of θ we have $\cos\theta > 0$, and hence

$$\sqrt{a^2-x^2} = a\cos\theta.$$

10.3.1. *Example.* To find

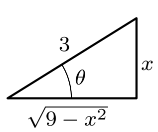
$$J = \int \sqrt{9-x^2}\,\mathrm{d}x$$

we substitute $x = 3\sin\theta$. Since θ ranges between $-\frac{\pi}{2}$ and $+\frac{\pi}{2}$ we have $\cos\theta > 0$ and thus

$$\sqrt{9-x^2} = \sqrt{9-9\sin^2\theta} = 3\sqrt{1-\sin^2\theta} = 3\sqrt{\cos^2\theta} = 3|\cos\theta| = 3\cos\theta.$$

We also have $\mathrm{d}x = 3\cos\theta\,\mathrm{d}\theta$, which then leads to

$$J = \int 3\cos\theta\, 3\cos\theta\,\mathrm{d}\theta = 9\int \cos^2\theta\,\mathrm{d}\theta.$$



$x = 3\sin\theta$
$\sqrt{9-x^2} = 3\cos\theta.$

This example shows that the integral we get after a trigonometric substitution is not always easy and may still require more tricks to be computed. For this particular integral we use the "double angle trick." Just as in § 3 we find

$$J = 9\int \cos^2\theta\,\mathrm{d}\theta = \frac{9}{2}\left(\theta + \tfrac{1}{2}\sin 2\theta\right) + C.$$

The last step is to undo the substitution $x = 3\sin\theta$. There are several strategies: one approach is to get rid of the double angles again and write all trigonometric expressions in terms of $\sin\theta$ and $\cos\theta$.

Since θ ranges between $-\frac{\pi}{2}$ and $+\frac{\pi}{2}$ we have

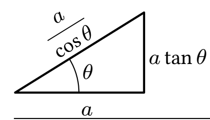
$$x = 3\sin\theta \iff \theta = \arcsin\frac{x}{3},$$

To substitute $\theta = \arcsin(\cdots)$ in $\sin 2\theta$ we need a double angle formula,

$$\sin 2\theta = 2\sin\theta\cos\theta = 2\cdot\frac{x}{3}\cdot\frac{\sqrt{9-x^2}}{3} = \frac{2}{9}x\sqrt{9-x^2}.$$

We get

$$\int\sqrt{9-x^2}\,\mathrm{d}x = \frac{9}{2}\theta + \frac{9}{2}\sin\theta\cos\theta + C. \qquad = \frac{9}{2}\arcsin\frac{x}{3} + \frac{1}{2}x\sqrt{9-x^2} + C.$$



$x = a\tan\theta$

$\sqrt{a^2+x^2} = a/\cos\theta$

$y = a/\cos\theta$

$\sqrt{y^2-a^2} = a\tan\theta$

10.4. Integrals containing $\sqrt{x^2-a^2}$ or $\sqrt{a^2+x^2}$. There are trigonometric substitutions that will remove either $\sqrt{x^2-a^2}$ or $\sqrt{a^2+x^2}$ from an integral. In both cases they come from the identities

$$\left(\frac{1}{\cos\theta}\right)^2 = \tan^2\theta + 1 \quad\text{or}\quad \left(\frac{1}{\cos\theta}\right)^2 - 1 = \tan^2\theta. \tag{9}$$

You can remember these identities either by drawing a right triangle with angle θ and with base of length 1, or else by dividing both sides of the equations

$$1 = \sin^2\theta + \cos^2\theta \quad\text{or}\quad 1-\cos^2\theta = \sin^2\theta$$

by $\cos^2\theta$.

10.4.1. *Example – turn the integral $\int_2^4\sqrt{x^2-4}\,\mathrm{d}x$ into a trigonometric integral.* Since $\sqrt{x^2-4} = \sqrt{x^2-2^2}$ we substitute

$$x = \frac{2}{\cos\theta},$$

which then leads to

$$\sqrt{x^2-4} = \sqrt{4\tan^2\theta} = 2\tan\theta.$$

In this last step we have to be careful with the sign of the square root: since $2 < x < 4$ in our integral, we can assume that $0 < \theta < \frac{\pi}{2}$ and thus that $\tan\theta > 0$. Therefore $\sqrt{\tan^2\theta} = \tan\theta$ instead of $-\tan\theta$.

The substitution $x = \frac{2}{\cos\theta}$ also implies that

$$\mathrm{d}x = 2\frac{\sin\theta}{\cos^2\theta}\,\mathrm{d}\theta.$$

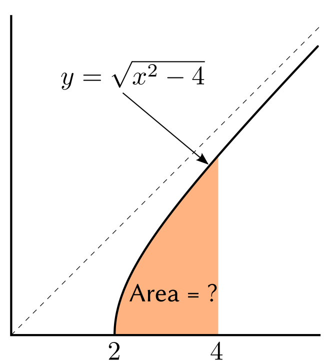


Figure 1. What is the area of the shaded region under the hyperbola? We first try to compute it using a trigonometric substitution (§ 10.4.1), and then using a rational substitution involving the U and V functions (§ 11.1.1). The answer turns out to be $4\sqrt{3} - 2\ln(2+\sqrt{3})$.

We finally also consider the integration bounds:

$$x = 2 \implies \frac{2}{\cos\theta} = 2 \implies \cos\theta = 1 \implies \theta = 0,$$

and

$$x = 4 \implies \frac{2}{\cos\theta} = 4 \implies \cos\theta = \tfrac{1}{2} \implies \theta = \frac{\pi}{3}.$$

Therefore we have

$$\int_2^4 \sqrt{x^2-4}\,\mathrm{d}x = \int_0^{\pi/3} 2\tan\theta \cdot 2\frac{\sin\theta}{\cos^2\theta}\,\mathrm{d}\theta = 4\int_0^{\pi/3} \frac{\sin^2\theta}{\cos^3\theta}\,\mathrm{d}\theta.$$

This integral is still not easy: it can be done by integration by parts, and you have to know the antiderivative of $1/\cos\theta$.

11. Rational substitution for integrals containing $\sqrt{x^2-a^2}$ or $\sqrt{a^2+x^2}$

11.1. The functions $U(t)$ and $V(t)$. Instead of using a trigonometric substitution one can also use the following identity to get rid of either $\sqrt{x^2-a^2}$ or $\sqrt{x^2+a^2}$. The identity is a relation between two functions U and V of a new variable t defined by

$$U(t) = \tfrac{1}{2}\Big(t + \frac{1}{t}\Big), \qquad V(t) = \tfrac{1}{2}\Big(t - \frac{1}{t}\Big). \tag{10}$$

These satisfy

$$U^2 = V^2 + 1, \tag{11}$$

which one can verify by direct substitution of the definitions (10) of $U(t)$ and $V(t)$.

To undo the substitution it is useful to note that if U and V are given by (10), then

$$t = U + V, \qquad \frac{1}{t} = U - V. \tag{12}$$

11.1.1. *Example § 10.4.1 again.* Here we compute the integral

$$A = \int_2^4 \sqrt{x^2-4}\,\mathrm{d}x$$

using the rational substitution (10).

Since the integral contains the expression $\sqrt{x^2-4} = \sqrt{x^2-2^2}$ we substitute $x = 2U(t)$. Using $U^2 = 1 + V^2$ we then have

$$\sqrt{x^2-4} = \sqrt{4U(t)^2-4} = 2\sqrt{U(t)^2-1} = 2|V(t)|.$$

When we use the substitution $x = aU(t)$ we should always assume that $t \geq 1$. Under that assumption we have $V(t) \geq 0$ (see Figure 2) and therefore $\sqrt{x^2-4} = 2V(t)$. To summarize, we have

$$x = 2U(t), \qquad \sqrt{x^2-4} = 2V(t). \tag{13}$$

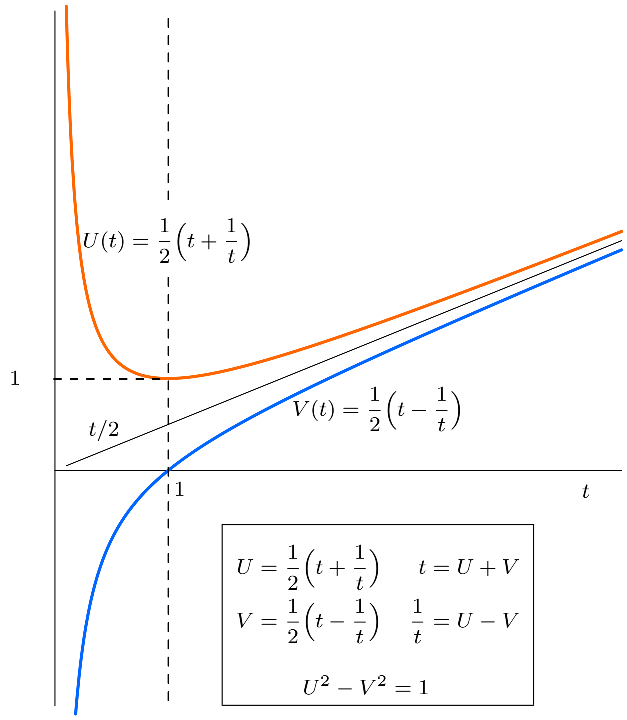


Figure 2. The functions $U(t)$ and $V(t)$

We can now do the indefinite integral:

$$\begin{aligned}
\int \sqrt{x^2-4}\,\mathrm{d}x &= \int \underbrace{2V(t)}_{\sqrt{x^2-4}} \cdot \underbrace{2U'(t)\,\mathrm{d}t}_{\mathrm{d}x} \\
&= \int 2 \cdot \frac{1}{2}\left(t - \frac{1}{t}\right) \cdot \left(1 - \frac{1}{t^2}\right)\,\mathrm{d}t \\
&= \int \left(t - \frac{2}{t} + \frac{1}{t^3}\right)\mathrm{d}t \\
&= \frac{t^2}{2} - 2\ln t - \frac{1}{2t^2} + C
\end{aligned}$$

To finish the computation we still have to convert back to the original x variable, and substitute the integration bounds. The most straightforward approach is to substitute $t = U + V$, and then remember the relations (13) between U, V, and x. Using these relations the middle term in the integral we just found becomes

$$-2\ln t = -2\ln(U+V) = -2\ln\left\{\frac{x}{2} + \sqrt{\left(\frac{x}{2}\right)^2 - 1}\right\}.$$

We can save ourselves some work by taking the other two terms together and factoring them as follows

$$
\begin{aligned}
\frac{t^2}{2} - \frac{1}{2t^2} &= \frac{1}{2}\Big(t^2 - \big(\frac{1}{t}\big)^2\Big) && a^2 - b^2 = (a+b)(a-b) \\
&= \frac{1}{2}\Big(t + \frac{1}{t}\Big)\Big(t - \frac{1}{t}\Big) && t + \tfrac{1}{t} = x \\
&= \frac{1}{2}x \cdot 2\sqrt{\big(\frac{x}{2}\big)^2 - 1} && \tfrac{1}{2}\big(t - \tfrac{1}{t}\big) = \sqrt{(\tfrac{x}{2})^2 - 1} \\
&= \frac{x}{2}\sqrt{x^2 - 4}.
\end{aligned} \tag{14}
$$

So we find

$$\int \sqrt{x^2 - 4}\ \mathrm{d}x = \frac{x}{2}\sqrt{x^2 - 4} - 2\ln\Big\{\frac{x}{2} + \sqrt{\big(\frac{x}{2}\big)^2 - 1}\Big\} + C.$$

Hence, substituting the integration bounds $x = 2$ and $x = 4$, we get

$$
\begin{aligned}
A &= \int_2^4 \sqrt{x^2 - 4}\ \mathrm{d}x \\
&= \Big[\frac{x}{2}\sqrt{x^2 - 4} - 2\ln\Big\{\frac{x}{2} + \sqrt{\big(\frac{x}{2}\big)^2 - 1}\Big\}\Big]_{x=2}^{x=4} \\
&= \frac{4}{2}\sqrt{16 - 4} - 2\ln\big(2 + \sqrt{3}\big) && \left(\begin{array}{c}\text{the terms with} \\ x = 2 \text{ vanish}\end{array}\right) \\
&= 4\sqrt{3} - 2\ln\big(2 + \sqrt{3}\big).
\end{aligned}
$$

11.1.2. *An example with* $\sqrt{1 + x^2}$. There are several ways to compute

$$I = \int \sqrt{1 + x^2}\, \mathrm{d}x$$

and unfortunately none of them are very simple. The simplest solution is to avoid finding the integral and look it up in a table, such as Table 2. But how were the integrals in that table found? One approach is to use the same pair of functions $U(t)$ and $V(t)$ from (10). Since $U^2 = 1 + V^2$ the substitution $x = V(t)$ allows us to take the square root of $1 + x^2$, namely,

$$x = V(t) \implies \sqrt{1 + x^2} = U(t).$$

Also, $\mathrm{d}x = V'(t)\mathrm{d}t = \frac{1}{2}\big(1 + \frac{1}{t^2}\big)\mathrm{d}t$, and thus we have

$$
\begin{aligned}
I &= \int \underbrace{\sqrt{1 + x^2}}_{=U(t)}\ \underbrace{\mathrm{d}x}_{\mathrm{d}V(t)} \\
&= \int \frac{1}{2}\big(t + \frac{1}{t}\big)\frac{1}{2}\big(1 + \frac{1}{t^2}\big)\ \mathrm{d}t \\
&= \frac{1}{4}\int \big(t + \frac{2}{t} + \frac{1}{t^3}\big)\ \mathrm{d}t \\
&= \frac{1}{4}\Big\{\frac{t^2}{2} + 2\ln t - \frac{1}{2t^2}\Big\} + C \\
&= \frac{1}{8}\big(t^2 - \frac{1}{t^2}\big) + \frac{1}{2}\ln t + C.
\end{aligned}
$$

At this point we have done the integral, but we should still rewrite the result in terms of the original variable x. We could use the same algebra as in (14), but this is not the only possible approach. Instead we could also use the relations (12), i.e.

$$t = U + V \text{ and } \frac{1}{t} = U - V$$

These imply

$$\begin{array}{rcccl} t^2 &=& (U+V)^2 &=& U^2 + 2UV + V^2 \\ t^{-2} &=& (U-V)^2 &=& U^2 - 2UV + V^2 \\ \hline t^2 - t^{-2} &=& \cdots &=& 4UV \end{array}$$

and conclude

$$\begin{aligned} I &= \int \sqrt{1+x^2}\,\mathrm{d}x \\ &= \frac{1}{8}\left(t^2 - \frac{1}{t^2}\right) + \frac{1}{2}\ln t + C \\ &= \frac{1}{2}UV + \frac{1}{2}\ln(U+V) + C \\ &= \frac{1}{2}x\sqrt{1+x^2} + \frac{1}{2}\ln\bigl(x + \sqrt{1+x^2}\bigr) + C. \end{aligned}$$

12. Simplifying $\sqrt{ax^2+bx+c}$ by completing the square

Any integral involving an expression of the form $\sqrt{ax^2+bx+c}$ can be reduced by means of a substitution to one containing one of the three forms $\sqrt{1-u^2}$, $\sqrt{u^2-1}$, or $\sqrt{u^2+1}$. We can achieve this reduction by completing the square of the quadratic expression under the square root. Once the more complicated square root $\sqrt{ax^2+bx+c}$ has been simplified to $\sqrt{\pm u^2 \pm 1}$, we can use either a trigonometric substitution, or the rational substitution from the previous section. In some cases the end result is one of the integrals listed in Table 2:

$$\int \frac{\mathrm{d}u}{\sqrt{1-u^2}} = \arcsin u \qquad \int \sqrt{1-u^2}\,\mathrm{d}u = \tfrac{1}{2}u\sqrt{1-u^2} + \tfrac{1}{2}\arcsin u$$

$$\int \frac{\mathrm{d}u}{\sqrt{1+u^2}} = \ln\bigl(u + \sqrt{1+u^2}\bigr) \qquad \int \sqrt{1+u^2}\,\mathrm{d}u = \tfrac{1}{2}u\sqrt{1+u^2} + \tfrac{1}{2}\ln\bigl(u + \sqrt{1+u^2}\bigr)$$

$$\int \frac{\mathrm{d}u}{\sqrt{u^2-1}} = \ln\bigl(u + \sqrt{u^2-1}\bigr) \qquad \int \sqrt{u^2-1}\,\mathrm{d}u = \tfrac{1}{2}u\sqrt{u^2-1} - \tfrac{1}{2}\ln\bigl(u + \sqrt{u^2-1}\bigr)$$

(all integrals "$+C$")

Table 2. Useful integrals. Except for the first one these should *not* be memorized.

Here are three examples. The problems have more examples.

12.1. Example. Compute

$$I = \int \frac{\mathrm{d}x}{\sqrt{6x - x^2}}.$$

Notice that since this integral contains a square root the variable x may not be allowed to have all values. In fact, the quantity $6x - x^2 = x(6-x)$ under the square root has to be positive so x must lie between $x = 0$ and $x = 6$. We now complete the square:

$$\begin{aligned} 6x - x^2 &= -\bigl(x^2 - 6x\bigr) \\ &= -\bigl(x^2 - 6x + 9 - 9\bigr) \\ &= -\bigl[(x-3)^2 - 9\bigr] \\ &= -9\Bigl[\frac{(x-3)^2}{9} - 1\Bigr] \\ &= -9\Bigl[\Bigl(\frac{x-3}{3}\Bigr)^2 - 1\Bigr]. \end{aligned}$$

At this point we decide to substitute

$$u = \frac{x-3}{3},$$

which leads to

$$\sqrt{6x - x^2} = \sqrt{-9\bigl(u^2 - 1\bigr)} = \sqrt{9\bigl(1 - u^2\bigr)} = 3\sqrt{1-u^2},$$
$$x = 3u + 3, \qquad \mathrm{d}x = 3\,\mathrm{d}u.$$

Applying this change of variable to the integral we get

$$\int \frac{\mathrm{d}x}{\sqrt{6x - x^2}} = \int \frac{3\mathrm{d}u}{3\sqrt{1-u^2}} = \int \frac{\mathrm{d}u}{\sqrt{1-u^2}} = \arcsin u + C = \arcsin \frac{x-3}{3} + C.$$

12.2. Example. Compute

$$I = \int \sqrt{4x^2 + 8x + 8}\,\mathrm{d}x.$$

We again complete the square in the quadratic expression under the square root:

$$4x^2 + 8x + 8 = 4\bigl(x^2 + 2x + 2\bigr) = 4\bigl\{(x+1)^2 + 1\bigr\}.$$

Thus we substitute $u = x + 1$, which implies $\mathrm{d}u = \mathrm{d}x$, after which we find

$$I = \int \sqrt{4x^2 + 8x + 8}\,\mathrm{d}x = \int 2\sqrt{(x+1)^2 + 1}\,\mathrm{d}x = 2\int \sqrt{u^2 + 1}\,\mathrm{d}u.$$

This last integral is in table 2, so we have

$$\begin{aligned} I &= u\sqrt{u^2+1} + \ln\bigl(u + \sqrt{u^2+1}\bigr) + C \\ &= (x+1)\sqrt{(x+1)^2+1} + \ln\Bigl\{x + 1 + \sqrt{(x+1)^2+1}\Bigr\} + C. \end{aligned}$$

12.3. Example. Compute:

$$I = \int \sqrt{x^2 - 4x - 5}\,\mathrm{d}x.$$

We first complete the square

$$\begin{aligned} x^2 - 4x - 5 &= x^2 - 4x + 4 - 9 \\ &= (x-2)^2 - 9 && u^2 - a^2 \text{ form} \\ &= 9\Bigl\{\bigl(\frac{x-2}{3}\bigr)^2 - 1\Bigr\} && u^2 - 1 \text{ form} \end{aligned}$$

This prompts us to substitute

$$u = \frac{x-2}{3}, \qquad \mathrm{d}u = \tfrac{1}{3}\mathrm{d}x,\ \text{i.e. } \mathrm{d}x = 3\,\mathrm{d}u.$$

We get

$$I = \int \sqrt{9\Big\{\big(\frac{x-2}{3}\big)^2 - 1\Big\}}\,\mathrm{d}x = \int 3\sqrt{u^2-1}\,3\,\mathrm{d}u = 9\int \sqrt{u^2-1}\,\mathrm{d}u.$$

Using the integrals in Table 2 and then undoing the substitution we find

$$\begin{aligned}
I &= \int \sqrt{x^2-4x-5}\,\mathrm{d}x \\
&= \tfrac{9}{2}u\sqrt{u^2-1} - \tfrac{9}{2}\ln\big(u+\sqrt{u^2-1}\big) + C \\
&= \tfrac{9}{2}\frac{x-2}{3}\sqrt{\big(\frac{x-2}{3}\big)^2-1} - \tfrac{9}{2}\ln\Big\{\frac{x-2}{3} + \sqrt{\big(\frac{x-2}{3}\big)^2-1}\Big\} + C \\
&= \tfrac{1}{2}(x-2)\sqrt{(x-2)^2-9} - \tfrac{9}{2}\ln\tfrac{1}{3}\Big\{x-2+\sqrt{(x-2)^2-9}\Big\} + C \\
&= \tfrac{1}{2}(x-2)\sqrt{x^2-4x+5} - \tfrac{9}{2}\ln\big\{x-2+\sqrt{x^2-4x+5}\big\} - \tfrac{9}{2}\ln\tfrac{1}{3} + C \\
&= \tfrac{1}{2}(x-2)\sqrt{x^2-4x+5} - \tfrac{9}{2}\ln\big\{x-2+\sqrt{x^2-4x+5}\big\} + \tilde{C}
\end{aligned}$$

13. Problems

Evaluate these integrals:

In any of these integrals, a is a positive constant.

1. $\displaystyle\int \frac{\mathrm{d}x}{\sqrt{1-x^2}}$

2. $\displaystyle\int \frac{\mathrm{d}x}{\sqrt{4-x^2}}$

3. $\displaystyle\int \sqrt{1+x^2}\,\mathrm{d}x$

4. $\displaystyle\int \frac{\mathrm{d}x}{\sqrt{2x-x^2}}$

5. $\displaystyle\int \frac{x\,\mathrm{d}x}{\sqrt{1-4x^4}}$

6. $\displaystyle\int_{-1/2}^{1/2} \frac{\mathrm{d}x}{\sqrt{4-x^2}}$

7. $\displaystyle\int_{-1}^{1} \frac{\mathrm{d}x}{\sqrt{4-x^2}}$

8. $\displaystyle\int_{0}^{\sqrt{3}/2} \frac{\mathrm{d}x}{\sqrt{1-x^2}}$

9. $\displaystyle\int \frac{\mathrm{d}x}{x^2+1}$

10. $\displaystyle\int \frac{\mathrm{d}x}{x^2+a^2}$

11. $\displaystyle\int \frac{\mathrm{d}x}{7+3x^2}$

12. $\displaystyle\int \frac{1+x}{a+x^2}\mathrm{d}x$

13. $\displaystyle\int \frac{\mathrm{d}x}{3x^2+6x+6}$

14. $\displaystyle\int \frac{\mathrm{d}x}{3x^2+6x+15}$

15. $\displaystyle\int_{1}^{\sqrt{3}} \frac{\mathrm{d}x}{x^2+1},$

16. $\displaystyle\int_{a}^{a\sqrt{3}} \frac{\mathrm{d}x}{x^2+a^2}.$

14. Chapter summary

There are several methods for finding the antiderivative of a function. Each of these methods allow us to transform a given integral into another, hopefully simpler, integral. Here are the methods that were presented in this chapter, in the order in which they appeared:

(1) *Double angle formulas and other trig identities*: some integrals can be simplified by using a trigonometric identity. This is not a general method, and only works for certain very specific integrals. Since these integrals do come up with some frequency it is worth knowing the double angle trick and its variations.

(2) *Integration by parts*: a very general formula; repeated integration by parts is done using reduction formulas.
(3) *Partial Fraction Decomposition*: a method that allows us to integrate any rational function.
(4) *Trigonometric and Rational Substitution*: a specific group of substitutions that can be used to simplify integrals containing the expression $\sqrt{ax^2+bx+c}$.

15. Mixed Integration Problems

One of the challenges in integrating a function is to recognize which of the methods we know will be most useful – so here is an unsorted list of integrals for practice.

Evaluate these integrals:

1. $\displaystyle\int_0^a x\sin x\,\mathrm{d}x$ •

2. $\displaystyle\int_0^a x^2\cos x\,\mathrm{d}x$ •

3. $\displaystyle\int_3^4 \frac{x\,\mathrm{d}x}{\sqrt{x^2-1}}$ •

4. $\displaystyle\int_{1/4}^{1/3} \frac{x\,\mathrm{d}x}{\sqrt{1-x^2}}$ •

5. $\displaystyle\int_3^4 \frac{\mathrm{d}x}{x\sqrt{x^2-1}}$ •

6. $\displaystyle\int \frac{x\,\mathrm{d}x}{x^2+2x+17}$ •

6. $\displaystyle\int \frac{x\,\mathrm{d}x}{\sqrt{x^2+2x+17}}$ •

7. $\displaystyle\int \frac{x^4}{(x^2-36)^{1/2}}\,\mathrm{d}x$

8. $\displaystyle\int \frac{x^4}{x^2-36}\,\mathrm{d}x$

9. $\displaystyle\int \frac{x^4}{36-x^2}\,\mathrm{d}x$

10. $\displaystyle\int \frac{x^2+1}{x^4-x^2}\,\mathrm{d}x$ •

11. $\displaystyle\int \frac{x^2+3}{x^4-2x^2}\,\mathrm{d}x$

12. $\displaystyle\int \frac{\mathrm{d}x}{(x^2-3)^{1/2}}$

13. $\displaystyle\int e^x(x+\cos(x))\,\mathrm{d}x$

14. $\displaystyle\int (e^x+\ln(x))\,\mathrm{d}x$

15. $\displaystyle\int \frac{\mathrm{d}x}{(x+5)\sqrt{x^2+5x}}$ hint: $x+5=1/u$

16. $\displaystyle\int \frac{x\,\mathrm{d}x}{x^2+2x+17}$

17. $\displaystyle\int \frac{x\,\mathrm{d}x}{x^2+2x+17}$

18. $\displaystyle\int \frac{x\,\mathrm{d}x}{x^2+2x+17}$

19. $\displaystyle\int \frac{x\,\mathrm{d}x}{x^2+2x+17}$

20. $\displaystyle\int \frac{3x^2+2x-2}{x^3-1}\,\mathrm{d}x$

21. $\displaystyle\int \frac{x^4}{x^4-16}\,\mathrm{d}x$

22. $\displaystyle\int \frac{x}{(x-1)^3}\,\mathrm{d}x$

23. $\displaystyle\int \frac{4}{(x-1)^3(x+1)}\,\mathrm{d}x$

24. $\displaystyle\int \frac{1}{\sqrt{6-2x-4x^2}}\,\mathrm{d}x$

25. $\displaystyle\int \frac{\mathrm{d}x}{\sqrt{x^2+2x+3}}$

26. $\displaystyle\int_1^e x\ln x\,\mathrm{d}x$

27. $\displaystyle\int 2x\ln(x+1)\,\mathrm{d}x$ •

28. $\displaystyle\int_{e^2}^{e^3} x^2\ln x\,\mathrm{d}x$

29. $\displaystyle\int_1^e x(\ln x)^3\,\mathrm{d}x$

30. $\displaystyle\int \arctan(\sqrt{x})\,\mathrm{d}x$ •

31. $\displaystyle\int x(\cos x)^2\,\mathrm{d}x$

32. $\displaystyle\int_0^\pi \sqrt{1+\cos(6w)}\,\mathrm{d}w$

Hint: $1+\cos\alpha = 2\sin^2\frac{\alpha}{2}$.

33. $\displaystyle\int \frac{1}{1+\sin(x)}\,\mathrm{d}x$

34. Find

$$\int \frac{\mathrm{d}x}{x(x-1)(x-2)(x-3)}$$

and

$$\int \frac{(x^3+1)\,\mathrm{d}x}{x(x-1)(x-2)(x-3)}$$

35. Compute

$$\int \frac{\mathrm{d}x}{x^3+x^2+x+1}$$

(Hint: to factor the denominator begin with $1+x+x^2+x^3 = (1+x)+x^2(1+x) = \dots$)

36. **[Group Problem]** You don't always have to find the antiderivative to find a definite integral. This problem gives you two examples of how you can avoid finding the antiderivative.

(a) To find

$$I = \int_0^{\pi/2} \frac{\sin x\,\mathrm{d}x}{\sin x + \cos x}$$

you use the substitution $u = \pi/2 - x$. The new integral you get must of course be equal to the integral I you started with, so if you *add the old and new integrals* you get $2I$. If you actually do this you will see that the sum of the old and new integrals is *very* easy to compute.

(b) Use your answer from **(a)** to compute

$$\int_0^1 \frac{\mathrm{d}x}{x+\sqrt{1-x^2}}.$$

(c) Use the same trick to find $\displaystyle\int_0^{\pi/2} \sin^2 x\,\mathrm{d}x$

37. **The Astroid.** Draw the curve whose equation is

$$|x|^{\frac{2}{3}} + |y|^{\frac{2}{3}} = a^{\frac{2}{3}},$$

where a is a positive constant. The curve you get is called the *Astroid*. Compute the area bounded by this curve.

38. **The Bow-Tie Graph.** Draw the curve given by the equation

$$y^2 = x^4 - x^6.$$

Compute the area bounded by this curve.

39. **The Fan-Tailed Fish.** Draw the curve given by the equation

$$y^2 = x^2\left(\frac{1-x}{1+x}\right).$$

Find the area enclosed by the loop. (Hint: after finding which integral you need to compute, substitute $x = \sin\theta$ to do the integral.)

40. Find the area of the region bounded by the curves

$$x = 2, \qquad y = 0, \qquad y = x\ln\frac{x}{2}$$

41. Find the volume of the solid of revolution obtained by rotating around the $x-$axis the region bounded by the lines $x = 5$, $x = 10$, $y = 0$, and the curve

$$y = \frac{x}{\sqrt{x^2+25}}.$$

42. *How to find the integral of* $f(x) = \dfrac{1}{\cos x}$.

Note that

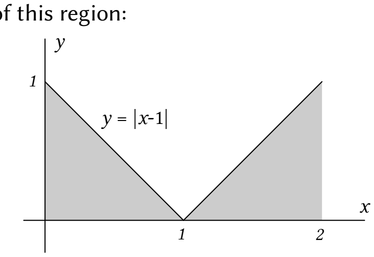
$$\frac{1}{\cos x} = \frac{\cos x}{\cos^2 x} = \frac{\cos x}{1-\sin^2 x},$$

and apply the substitution $s = \sin x$ followed by a partial fraction decomposition to compute $\displaystyle\int \frac{\mathrm{d}x}{\cos x}$.

Calculus bloopers

As you'll see, the following computations can't be right; but where did they go wrong?

43. Here is a failed computation of the area of this region:



Clearly the combined area of the two triangles should be 1. Now let's try to get this answer by integration.

Consider $\int_0^2 |x-1|\,\mathrm{d}x$. Let $f(x) = |x-1|$ so that

$$f(x) = \begin{cases} x-1 & \text{if } x \geq 1 \\ 1-x & \text{if } x < 1 \end{cases}$$

Define

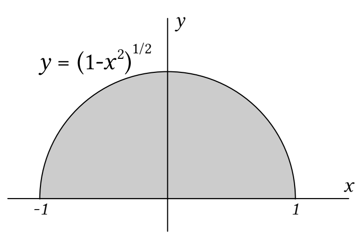
$$F(x) = \begin{cases} \frac{1}{2}x^2 - x & \text{if } x \geq 1 \\ x - \frac{1}{2}x^2 & \text{if } x < 1 \end{cases}$$

Then since F is an antiderivative of f we have by the Fundamental Theorem of Calculus:

$$\begin{aligned} \int_0^2 |x-1|\,\mathrm{d}x &= \int_0^2 f(x)\,\mathrm{d}x \\ &= F(2) - F(0) \\ &= \Big(\frac{2^2}{2} - 2\Big) - \Big(0 - \frac{0^2}{2}\Big) \\ &= 0. \end{aligned}$$

But this integral cannot be zero, $f(x)$ is positive except at one point. How can this be?

44. *It turns out that the area enclosed by a circle is zero – ?!* According to the ancient formula for the area of a disc, the area of the following half-disc is $\pi/2$.



We can also compute this area by means of an integral, namely

$$\text{Area} = \int_{-1}^{1} \sqrt{1-x^2}\,\mathrm{d}x$$

Substitute $u = 1 - x^2$ so:

$$u = 1 - x^2, \quad x = \sqrt{1-u} = (1-u)^{\frac{1}{2}},$$
$$\mathrm{d}x = (\frac{1}{2})(1-u)^{-\frac{1}{2}}(-1)\,\mathrm{d}u.$$

Hence

$$\int \sqrt{1-x^2}\,\mathrm{d}x = \int \sqrt{u}\,(\frac{1}{2})(1-u)^{-\frac{1}{2}}(-1)\,\mathrm{d}u.$$

Now take the definite integral from $x = -1$ to $x = 1$ and note that $u = 0$ when $x = -1$ and $u = 0$ also when $x = 1$, so

$$\int_{-1}^{1} \sqrt{1-x^2}\,\mathrm{d}x = \int_0^0 \sqrt{u}(\frac{1}{2})(1-u)^{-\frac{1}{2}}(-1)\,\mathrm{d}u = 0$$

The last being zero since

$$\int_0^0 (\text{ anything })\,\mathrm{d}x = 0.$$

But the integral on the left is equal to half the area of the unit disc. Therefore half a disc has zero area, and a whole disc should have twice as much area: still zero!

How can this be?

CHAPTER 2

Proper and Improper Integrals

All the definite integrals that we have seen so far were of the form

$$I = \int_a^b f(x)\,\mathrm{d}x,$$

where a and b are finite numbers, and where the integrand (the function $f(x)$) is "nice" on the interval $a \le x \le b$, i.e. the function $f(x)$ does not become infinite anywhere in the interval. There are many situations where one would like to compute an integral that fails one of these conditions; i.e. integrals where a or b is not finite, or where the integrand $f(x)$ becomes infinite somewhere in the interval $a \le x \le b$ (usually at an endpoint). Such integrals are called ***improper integrals***.

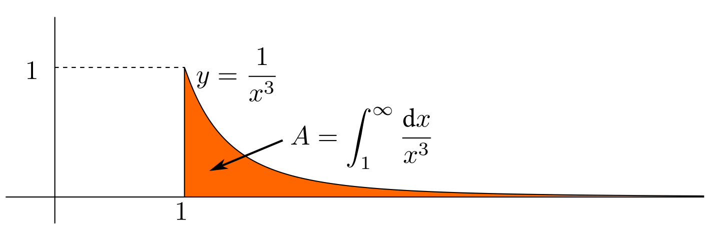
If we think of integrals as areas of regions in the plane, then improper integrals usually refer to areas of infinitely large regions so that some care must be taken in interpreting them. The formal definition of the integral as a limit of Riemann sums cannot be used since it assumes both that the integration bounds a and b are finite, and that the integrand $f(x)$ is bounded. Improper integrals have to be defined on a case by case basis. The next section shows the usual ways in which this is done.

1. Typical examples of improper integrals

1.1. Integral on an unbounded interval. Consider the integral

$$A = \int_1^\infty \frac{\mathrm{d}x}{x^3}.$$

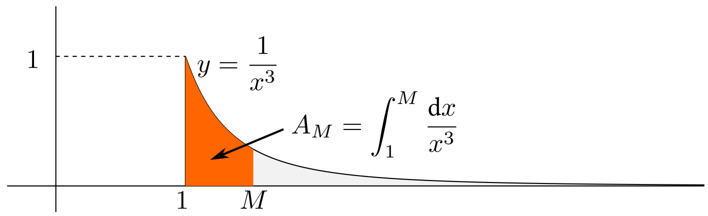
This integral has a new feature that we have not dealt with before, namely, one of the integration bounds is "∞" rather than a finite number. The interpretation of this integral is that it is the area of the region under the graph of $y = 1/x^3$, with $1 < x < \infty$.



Because the integral goes all the way to "$x = \infty$" the region whose area it represents stretches infinitely far to the right. Could such an infinitely wide region still have a finite area? And if it is, can we compute it? To compute the integral I that has the ∞ in its integration bounds, we first replace the integral by one that is more familiar, namely

$$A_M = \int_1^M \frac{\mathrm{d}x}{x^3},$$

where $M > 1$ is some finite number. This integral represents the area of a finite region, namely all points between the graph and the x-axis, and with $1 \le x \le M$.



We know how to compute this integral:

$$A_M = \int_1^M \frac{dx}{x^3} = \left[-\frac{1}{2x^2}\right]_1^M = -\frac{1}{2M^2} + \frac{1}{2}.$$

The area we find depends on M. The larger we choose M, the larger the region is and the larger the area should be. If we let $M \to \infty$ then the region under the graph between $x = 1$ and $x = M$ will expand and eventually fill up the whole region between graph and x-axis, and to the right of $x = 1$. Thus the area should be

$$A = \lim_{M\to\infty} A_M = \lim_{M\to\infty} \int_1^M \frac{dx}{x^3} = \lim_{M\to\infty} \left[-\frac{1}{2M^2} + \frac{1}{2}\right] = \frac{1}{2}.$$

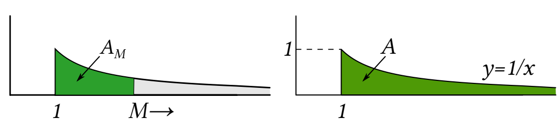
We conclude that the infinitely large region between the graph of $y = 1/x^3$ and the x-axis that lies to the right of the line $x = 1$ has finite area, and that this area is exactly $\frac{1}{2}$!

1.2. Second example on an unbounded interval.

The following integral is very similar to the one we just did:

$$A = \int_1^\infty \frac{dx}{x}.$$

The integral represents the area of the region that lies to the right of the line $x = 1$, and is caught between the x-axis and the hyperbola $y = 1/x$.



As in the previous example the region extends infinitely far to the right while at the same time becoming narrower and narrower. To see what its area is we again look at the truncated region that contains only those points between the graph and the x-axis, and for which $1 \le x \le M$. This area is

$$A_M = \int_1^M \frac{dx}{x} = \left[\ln x\right]_1^M = \ln M - \ln 1 = \ln M.$$

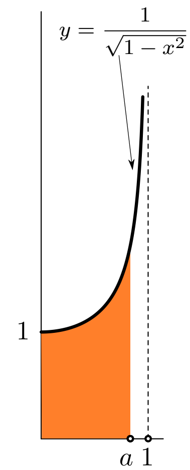
The area of the whole region with $1 \le x < \infty$ is the limit

$$A = \lim_{M\to\infty} A_M = \lim_{M\to\infty} \ln M = +\infty.$$

So we see that the area under the hyperbola is **not finite!**

1.3. An improper integral on a finite interval. In this third example we consider the integral

$$I = \int_0^1 \frac{\mathrm{d}x}{\sqrt{1-x^2}}.$$



The integration bounds for this integral are 0 and 1 so they are finite, but the integrand becomes infinite at one end of the integration interval:

$$\lim_{x \nearrow 1} \frac{1}{\sqrt{1-x^2}} = +\infty.$$

The region whose area the integral I represents does not extend infinitely far to the left or the right, but in this example it extends infinitely far upward. To compute this area we again truncate the region by looking at all points with $1 \leq x \leq a$ for some constant $a < 1$, and compute

$$I_a = \int_0^a \frac{\mathrm{d}x}{\sqrt{1-x^2}} = \big[\arcsin x\big]_0^a = \arcsin a.$$

The integral I is then the limit of I_a, i.e.

$$\int_0^1 \frac{\mathrm{d}x}{\sqrt{1-x^2}} = \lim_{a \nearrow 1} \int_0^a \frac{\mathrm{d}x}{\sqrt{1-x^2}} = \lim_{a \nearrow 1} \arcsin a = \arcsin 1 = \frac{\pi}{2}.$$

We see that the area is finite.

1.4. A doubly improper integral. Let us try to compute

$$I = \int_{-\infty}^{\infty} \frac{\mathrm{d}x}{1+x^2}.$$

This example has a new feature, namely, both integration limits are infinite. To compute this integral we replace them by finite numbers, i.e. we compute

$$\begin{aligned}
\int_{-\infty}^{\infty} \frac{\mathrm{d}x}{1+x^2} &= \lim_{a\to-\infty} \lim_{b\to\infty} \int_a^b \frac{\mathrm{d}x}{1+x^2} \\
&= \lim_{a\to-\infty} \lim_{b\to\infty} \arctan(b) - \arctan(a) \\
&= \lim_{b\to\infty} \arctan b - \lim_{a\to-\infty} \arctan a \\
&= \frac{\pi}{2} - \big(-\frac{\pi}{2}\big) \\
&= \pi.
\end{aligned}$$

A different way of getting the same example is to replace ∞ and $-\infty$ in the integral by a and $-a$ and then let $a \to \infty$. The only difference with our previous approach is that we now use one variable (a) instead of two (a and b). The computation goes as follows:

$$\begin{aligned}
\int_{-\infty}^{\infty} \frac{\mathrm{d}x}{1+x^2} &= \lim_{a\to\infty} \int_{-a}^{a} \frac{\mathrm{d}x}{1+x^2} \\
&= \lim_{a\to\infty} \Big(\arctan(a) - \arctan(-a)\Big) \\
&= \frac{\pi}{2} - \big(-\frac{\pi}{2}\big) = \pi.
\end{aligned}$$

In this example we got the same answer using either approach. This is not always the case, as the next example shows.

1.5. Another doubly improper integral. Suppose we try to compute the integral

$$I = \int_{-\infty}^{\infty} x \, \mathrm{d}x.$$

The shorter approach where we replace both $\pm\infty$ by $\pm a$ would be

$$\int_{-\infty}^{\infty} x \, \mathrm{d}x = \lim_{a\to\infty} \int_{-a}^{a} x \, \mathrm{d}x = \lim_{a\to\infty} \frac{a^2}{2} - \frac{(-a)^2}{2} = 0.$$

On the other hand, if we take the longer approach, where we replace $-\infty$ and ∞ by two different constants, then we get this

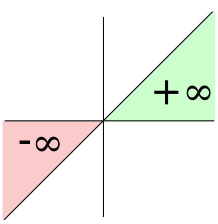
$$\begin{aligned} \int_{-\infty}^{\infty} x \, \mathrm{d}x &= \lim_{a\to-\infty} \lim_{b\to\infty} \int_{a}^{b} x \, \mathrm{d}x \\ &= \lim_{a\to-\infty} \lim_{b\to\infty} \left[\frac{x^2}{2}\right] \\ &= \lim_{b\to\infty} \frac{b^2}{2} - \lim_{a\to-\infty} \frac{a^2}{2}. \end{aligned}$$

At this point we see that both limits $\lim_{b\to\infty} b^2/2 = \infty$ and $\lim_{a\to-\infty} a^2/2 = \infty$ do not exist. The result we therefore get is

$$\int_{-\infty}^{\infty} x \, \mathrm{d}x = \infty - \infty.$$

Since ∞ is not a number we find that the improper integral does not exist.

We conclude that for some improper integrals different ways of computing them can give different results. This means that we have to be more precise and specify which definition of the integral we use. The next section lists the definitions that are commonly used.



$\int_{-\infty}^{\infty} x \, \mathrm{d}x$ is the area above minus the area below the axis.

2. Summary: how to compute an improper integral

2.1. How to compute an improper integral on an unbounded interval. By definition the improper integral

$$\int_{a}^{\infty} f(x) \, \mathrm{d}x$$

is given by

$$\int_{a}^{\infty} f(x) \, \mathrm{d}x = \lim_{M\to\infty} \int_{a}^{M} f(x) \, \mathrm{d}x. \tag{15}$$

This is how the integral in § 1.1 was computed.

2.2. How to compute an improper integral of an unbounded function. If the integration interval is bounded, but if the integrand becomes infinite at one of the endpoints, say at $x = a$, then we define

$$\int_{a}^{b} f(x) \, \mathrm{d}x = \lim_{s\searrow a} \int_{s}^{b} f(x) \, \mathrm{d}x. \tag{16}$$

2.3. Doubly improper integrals. If the integration interval is the whole real line, i.e. if we need to compute the integral

$$I = \int_{-\infty}^{\infty} f(x)\,\mathrm{d}x,$$

then we must replace both integration bound by finite numbers and then let those finite numbers go to $\pm\infty$. Thus we would first have to compute an antiderivative of f,

$$F(x) = \int f(x)\,\mathrm{d}x.$$

The Fundamental Theorem of Calculus then implies

$$I_{a,b} = \int_a^b f(x)\,\mathrm{d}x = F(b) - F(a)$$

and we set

$$I = \lim_{b\to\infty} F(b) - \lim_{a\to-\infty} F(a).$$

Note that according to this definition we have to compute both limits separately. The example in Section 1.5 shows that it really is necessary to do this, and that computing

$$\int_{-\infty}^{\infty} f(x)\mathrm{d}x = \lim_{a\to\infty} F(a) - F(-a)$$

can give different results.

In general, if an integral

$$\int_a^b f(x)\,\mathrm{d}x$$

is improper at both ends we must replace them by c, d with $a < c < d < b$ and compute the limit

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{c\searrow a}\lim_{d\nearrow b}\int_c^d f(x)\,\mathrm{d}x.$$

For instance,

$$\lim \int_{-1}^{1} \frac{\mathrm{d}x}{\sqrt{1-x^2}} = \lim_{c\searrow -1}\lim_{d\nearrow 1}\int_c^d \frac{\mathrm{d}x}{\sqrt{1-x^2}}\ldots$$

3. More examples

3.1. Area under an exponential. Let a be some positive constant and consider the graph of $y = e^{-ax}$ for $x > 0$. How much area is there between the graph and the x-axis with $x > 0$? (See Figure 1 for the cases $a = 1$ and $a = 2$.) The answer is given by the

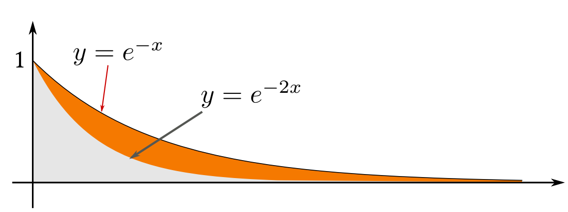


Figure 1. What is the area under the graph of $y = e^{-x}$? What fraction of the region under the graph of $y = e^{-x}$ lies under the graph of $y = e^{-2x}$?

improper integral

$$\begin{aligned}
A &= \int_0^\infty e^{-ax}\,\mathrm{d}x \\
&= \lim_{M\to\infty}\int_0^M e^{-ax}\,\mathrm{d}x \\
&= \lim_{M\to\infty}\left[-\frac{1}{a}e^{-ax}\right]_0^M \\
&= -\frac{1}{a}\lim_{M\to\infty}\Big(e^{-aM}-1\Big) && a>0 \text{ so } \lim_{M\to\infty} e^{-aM}=0 \\
&= \frac{1}{a}.
\end{aligned}$$

We see that the area under the graph is finite, and that it is given by $1/a$. In particular the area under the graph of $y=e^{-x}$ is exactly 1, while the area under the graph of $y=e^{-2x}$ is exactly half that (i.e. $1/2$).

3.2. Improper integrals involving x^{-p}. The examples in § 1.1 and § 1.2 are special cases of the following integral

$$I=\int_1^\infty \frac{\mathrm{d}x}{x^p}\ ,$$

where $p>0$ is some constant. We already know that the integral is $\frac{1}{2}$ if $p=3$ (§ 1.1), and also that the integral is infinite (does not exist) if $p=1$ (§ 1.2). We can compute it in the same way as before,

$$\begin{aligned}
I &= \lim_{M\to\infty}\int_1^M \frac{\mathrm{d}x}{x^p} \\
&= \lim_{M\to\infty}\left[\frac{1}{1-p}x^{1-p}\right]_1^M && \text{assume } p\neq 1 \\
&= \frac{1}{1-p}\lim_{M\to\infty}\Big(M^{1-p}-1\Big).
\end{aligned} \tag{17}$$

At this point we have to find out what happens to M^{1-p} as $M\to\infty$. This depends on the sign of the exponent $1-p$. If this exponent is positive, then M^{1-p} is a positive power of M and therefore becomes infinite as $M\to\infty$. On the other hand, if the exponent is negative then M^{1-p} is a negative power of M so that it goes to zero as $M\to\infty$. To summarize:

- If $0<p<1$ then $1-p>0$ so that $\lim_{M\to\infty} M^{1-p}=\infty$;
- if $p>1$ then $1-p<0$, and

$$\lim_{M\to\infty} M^{1-p}=\lim_{M\to\infty}\frac{1}{M^{p-1}}=0.$$

If we apply this to (17) then we find that

$$\text{if } 0<p\le 1 \text{ then } \int_1^\infty \frac{\mathrm{d}x}{x^p}=\infty,$$

$$\text{and if } p>1 \text{ then } \int_1^\infty \frac{\mathrm{d}x}{x^p}=\frac{1}{p-1}.$$

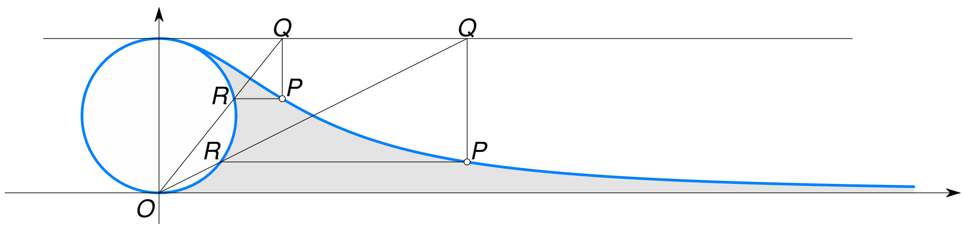


Figure 2. Geometric construction of the ***Versiera.*** The figure shows a circle of diameter 1 that touches the x-axis exactly at the origin. For any point Q on the line $y = 1$ we can define a new point P by intersecting the line segment OQ with the circle, and requiring P to have the same x-coordinate as Q and the same y-coordinate as R. The *Versiera* is the curve traced out by the point P as we lets the point Q slide along the line $y = 1$. See problem 4.16.

4. Problems

Compute the following improper integrals and draw the region whose area they represent:

1. $\displaystyle\int_0^\infty \frac{dx}{(2+x)^2}.$

2. $\displaystyle\int_0^{1/2} (2x-1)^{-3}\, dx$

3. $\displaystyle\int_0^3 \frac{dx}{\sqrt{3-x}}.$

4. $\displaystyle\int_0^1 1/\sqrt{x}\, dx.$

5. $\displaystyle\int_0^\infty xe^{-x^2}\, dx.$

6. $\displaystyle\int_1^\infty \frac{x-1}{x+x^2}\, dx.$

7. $\displaystyle\int_1^\infty \frac{x-1}{x^3+x^2}\, dx.$

8. $\displaystyle\int_0^7 \frac{dx}{\sqrt{x}}.$

9. $\displaystyle\int_0^1 \frac{dx}{x\sqrt{x}}.$

10. $\displaystyle\int_0^1 \frac{dx}{x+\sqrt{x}}.$ (suggestion: $x = u^2$)

11. $\displaystyle\int_0^1 \left(\frac{1}{\sqrt{x}} + \frac{1}{x}\right) dx.$

12. $\displaystyle\int_5^\infty e^{-2x}\, dx$

13. $\displaystyle\int_0^\infty xe^{-x}\, dx$

14. $\displaystyle\int_0^\infty xe^{-x^2}\, dx$

15. $\displaystyle\int_0^\infty e^{-\sqrt{x}}\, dx$

16. The graph of the function $y = 1/(1+x^2)$ is called the *Versiera* (see Figure 2). Compute the area of the shaded region between the Versiera and the circle in Figure 2.

17. Let a be a positive constant.

(a) Draw the graph of $y = xe^{-ax}$, $0 \le x < \infty$ (the function has one maximum: where is it?)

(b) Compute $I_a = \displaystyle\int_0^\infty xe^{-ax}\, dx$. How does I_a depend on a?

(c) Compute $\lim_{a\searrow 0} I_a$. Interpret your answer using the graphs of $y = x^{-ax}$ that you drew in part (i).

18. $\displaystyle\int_{-\infty}^0 e^x\, dx$. How would you define this integral (one of the integration bounds is $-\infty$ rather than $+\infty$)?

19. **(a)** $\displaystyle\int_0^\infty e^{-x} \sin \pi x\, dx$ where a, b are positive constants. (You have done the integral before – see Ch I, § 7, Problem 7.2 .)

(b) As in the previous problems, draw the region whose area the integral from (**a**) represents. Note that the function in the integral for this problem is not positive everywhere. How does this affect your answer?

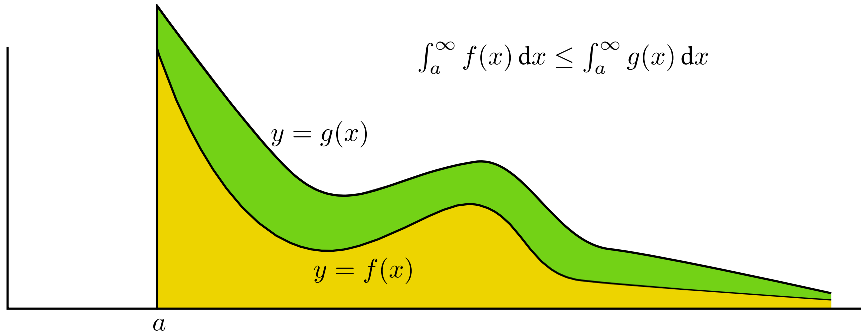


Figure 3. Comparing improper integrals. Here f and g are positive functions that satisfy $f(x) \leq g(x)$ for all $x \geq a$. If $\int_a^\infty g(x)\,\mathrm{d}x$ is finite, then so is $\int_a^\infty f(x)\,\mathrm{d}x$. Conversely, if $\int_a^\infty f(x)\,\mathrm{d}x$ is not finite, then $\int_a^\infty g(x)\,\mathrm{d}x$ cannot be finite either.

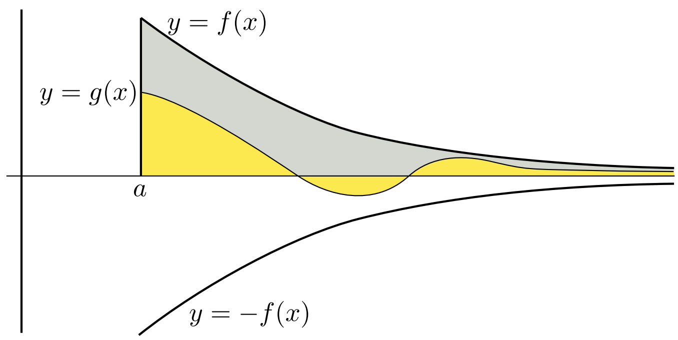


Figure 4. In this figure $f(x)$ is again positive, but $g(x)$ is bounded by $-f(x) \leq g(x) \leq f(x)$. The same conclusion still holds, namely, if $\int_a^\infty f(x)\mathrm{d}x$ exists, then so does $\int_a^\infty g(x)\mathrm{d}x$.

Can you tell from your drawing if the integral is positive?

20. (*A way to visualize that* $\int_1^\infty \frac{\mathrm{d}x}{x} = \infty$)

(a) Show that for any $a > 0$ one has

$$\int_a^{2a} \frac{\mathrm{d}x}{x} = \ln 2;$$

in particular, this integral is the same for all $a > 0$.

(b) Compute the area under the graph of $y = 1/x$ between $x = 1$ and $x = 2^n$ is $n \cdot \ln 2$ by splitting the region into:

- the part where $1 \leq x \leq 2$,
- the part where $2 \leq x \leq 4$,
- the part where $4 \leq x \leq 8$,

$\vdots$

- the part where $2^{n-1} \leq x \leq 2^n$.

(c) Explain how the answer to **(b)** implies that the integral $\int_1^\infty \mathrm{d}x/x$ does not exist.

21. The area under the graph of $y = 1/x$ with $1 \leq x < \infty$ is infinite. Compute the volume of the funnel-shaped solid you get by revolving this region around the x-axis. Is the volume of this funnel finite or infinite? •

5. Estimating improper integrals

Sometimes it is just not possible to compute an improper integral because we simply cannot find an antiderivative for the integrand. When this happens we can still try to estimate the integral by comparing it with easier integrals, and even if we cannot compute the integral we can still try to answer the question "does the integral exist," i.e.

$$\textit{Does } \lim_{M\to\infty} \int_a^M f(x)\,\mathrm{d}x \textit{ exist?}$$

In this section we will see a number of examples of improper integrals that much easier to estimate than to compute. Throughout there are three principles that we shall use:

Integral of a positive function. If the function $f(x)$ satisfies $f(x) \geq 0$ for all $x \geq a$ then either the integral $\int_a^\infty f(x)\mathrm{d}x$ exists, or else it is infinite, by which we mean

$$\lim_{M\to\infty} \int_a^M f(x)\,\mathrm{d}x = \infty.$$

If the function $f(x)$ is not always positive then the above limit can fail to exist by oscillating without ever becoming infinite. For an example see § 5.1.2; see also § 5.1 for further explanation.

Comparison with easier integrals. If $y = f(x)$ and $y = g(x)$ are functions defined for $x \geq a$, and if $|g(x)| \leq f(x)$ for all $x \geq a$, then

$$\left|\int_a^\infty g(x)\,\mathrm{d}x\right| \leq \int_a^\infty f(x)\,\mathrm{d}x.$$

In particular,

$$\int_a^\infty f(x)\,\mathrm{d}x \text{ exists} \implies \int_a^\infty g(x)\,\mathrm{d}x \text{ exists},$$
$$\int_a^\infty g(x)\,\mathrm{d}x \text{ does not exist} \implies \int_a^\infty f(x)\,\mathrm{d}x \text{ does not exist}.$$

Read § 5.2 for more details and examples.

Only the tail matters. If $f(x)$ is a continuous function for all $x \geq a$, then for any $b \geq a$ we have

$$\int_a^\infty f(x)\,\mathrm{d}x \text{ exists} \iff \int_b^\infty f(x)\,\mathrm{d}x \text{ exists}.$$

Furthermore, for any $b \geq a$ we have

$$\int_a^\infty f(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x + \int_b^\infty f(x)\,\mathrm{d}x. \tag{18}$$

This is further explained in Theorem 5.3 and the examples following it.

5.1. Improper integrals of positive functions. Suppose that the function $f(x)$ is defined for all $x \geq a$, and that $f(x) \geq 0$ for all x. To see if the improper integral $\int_a^\infty f(x)\,\mathrm{d}x$ exists we have to figure out what happens to

$$I_M = \int_a^M f(x)\,\mathrm{d}x$$

as $M \nearrow \infty$. Since we are assuming that $f(x) \geq 0$ the integral I_M represents the area under the graph of f up to $x = M$. As we let M increase this region expands, and thus the integral I_M increases. So, as $M \nearrow \infty$ there are two possibilities: either I_M remains finite and converges to a finite number, or I_M becomes infinitely large. The following theorem summarizes this useful observation:

5.1.1. *Theorem. If $f(x) \geq 0$ for all $x \geq a$ then either the integral*

$$I = \int_a^\infty f(x)\, dx$$

exists, (i.e. I is a finite number), or else

$$I = \int_a^\infty f(x)\, dx = \infty, \quad \textit{i.e.} \quad \lim_{M\to\infty} \int_a^M f(x)\, dx = \infty.$$

5.1.2. *Example - integral to infinity of the cosine.* To illustrate what Theorem 5.1.1 says, let's consider an improper integral of a function that is not always positive. For instance, consider

$$I = \int_0^\infty \cos x\ dx.$$

The function in the integral is $f(x) = \cos x$, and this function is clearly not always positive. When we try to compute this integral we get

$$I = \lim_{M\to\infty} \int_0^M \cos x\ dx = \lim_{M\to\infty} \Big[\sin x\Big]_{x=0}^M = \lim_{M\to\infty} \sin M.$$

This limit does not exist as $\sin M$ oscillates up and down between -1 and $+1$ as $M \to \infty$. On the other hand, since $\sin M$ stays between -1 and $+1$, we cannot say that

$$\lim_{M\to\infty} \int_0^M \cos x\ dx = +\infty.$$

Theorem 5.1.1 tells us that if we integrate a positive function then this kind of oscillatory behavior cannot occur.

5.2. Comparison Theorem for Improper Integrals. *Suppose $f(x)$ and $g(x)$ are functions that are defined for $a \leq x < \infty$, and suppose that $|g(x)| \leq f(x)$ for all $x \geq a$, i.e.*

$$-f(x) \leq g(x) \leq f(x) \textit{ for all } x \geq a.$$

If the improper integral $\int_a^\infty f(x)\, dx$ exists then the improper integral $\int_a^\infty g(x)\, dx$ also exists, and one has

$$\left| \int_a^\infty g(x)\ dx \right| \leq \int_a^\infty f(x)\ dx.$$

This theorem is used in two ways: it can be used to verify that some improper integral exists without actually computing the integral, and it can also be used to estimate an improper integral.

5.2.1. *Example.* Consider the improper integral

$$I = \int_1^\infty \frac{dx}{1+x^3}.$$

The function in the integral is a rational function so in principle we know how to compute the integral. It turns out the computation is not very easy. If we don't really need to know the exact value of the integral, but only want a rough estimate of the integral, then we could compare the integral with an easier integral.

To decide which simpler integral we should use as comparison, we reason as follows. Since "only the tail matters," we consider the integrand $\frac{1}{1+x^3}$ for large x. When x is very large x^3 will be much larger than 1, so that we may guess that we can ignore the "1" in the denominator $1 + x^3$:

$$\frac{1}{1+x^3} \approx \frac{1}{x^3} \quad \text{as } x \to \infty. \tag{19}$$

This suggests that we may be able to compare the given integral with the integral

$$\int_1^\infty \frac{1}{x^3}\,\mathrm{d}x.$$

We know from our very first example in this chapter (§ 1.1) that this last integral is finite (we found that it is $\frac{1}{2}$). Therefore we can guess that our integral I also is finite.

Now let's try to use the Comparison Theorem 5.2 to get certainty by proving that the integral I does indeed exist. We want to show that the integral $\int_1^\infty \frac{1}{1+x^3}\mathrm{d}x$ exists, so we choose

$$g(x) = \frac{1}{1+x^3}, \quad \text{and thus } f(x) = \frac{1}{x^3}.$$

We can compare these functions as follows:

$$\text{it follows from: } x^3 \le 1 + x^3 \text{ for all } x \ge 1 \qquad \left(\begin{array}{l}\text{divide both sides first}\\ \text{by } x^3 \text{ and then by } 1+\\ x^3\end{array}\right)$$

$$\text{that: } \frac{1}{1+x^3} \le \frac{1}{x^3} \text{ for } x \ge 1$$

This tells us that

$$\int_1^\infty \frac{\mathrm{d}x}{1+x^3} \le \int_1^\infty \frac{\mathrm{d}x}{x^3} = \frac{1}{2}.$$

Therefore we have found that the integral I does indeed exist, and that it is no more than $\frac{1}{2}$.

We can go beyond this and try to find a *lower bound* (instead of saying that I is no more than $\frac{1}{2}$ we try to say that it is at least as large as some other number.) Here is one way of doing that:

$$1 + x^3 \le x^3 + x^3 \text{ for all } x \ge 1$$

$$\implies 1 + x^3 \le 2x^3 \text{ for all } x \ge 1 \qquad \left(\begin{array}{l}\text{divide both sides first}\\ \text{by } 2x^3 \text{ and then by } 1+\\ x^3\end{array}\right)$$

$$\implies \frac{1}{2x^3} \le \frac{1}{1+x^3} \text{ for } x \ge 1$$

This implies that

$$\int_1^\infty \frac{\mathrm{d}x}{2x^3} \le \int_1^\infty \frac{\mathrm{d}x}{1+x^3}.$$

The first integral here is half the integral we computed in § 1.1, so we get

$$\frac{1}{4} \le \int_1^\infty \frac{\mathrm{d}x}{1+x^3}.$$

In summary, we have found that

$$\frac{1}{4} \le \int_1^\infty \frac{\mathrm{d}x}{1+x^3} \le \frac{1}{2}.$$

5.2.2. *Second example.* Does the integral

$$I = \int_1^\infty \frac{x}{x^2+1}\,\mathrm{d}x$$

exist? Since the function we are integrating is positive, we could also ask *is the integral finite?*

As in the previous example we look at the behavior of the integrand as $x \to \infty$. For large x we can assume that x^2 is much larger than x, so that it would be reasonable to ignore the 1 in the denominator $x^2 + 1$. If we do that than we find that

$$\frac{x}{x^2+1} \approx \frac{x}{x^2} = \frac{1}{x}.$$

If this were correct, then we would end up comparing the integral I with the simpler integral

$$\int_1^\infty \frac{1}{x}\,\mathrm{d}x.$$

We know this latter integral is not finite (it was our second example, see § 1.2) and therefore we guess that the integral I probably also is not finite. To give a sound argument we will use the Comparison Theorem.

Our goal is to show that the integral

$$I = \int_1^\infty f(x)\,\mathrm{d}x, \quad \text{with } f(x) = \frac{x}{1+x^2}$$

is not finite. To do this we have to find a function $g(x)$ such that

- $g(x)$ is smaller than $f(x)$ (so that $\int f\mathrm{d}x$ will be larger than $\int g(x)\mathrm{d}x$),
- $g(x)$ is easy to integrate, and
- the integral of $g(x)$ is not finite.

The first and last point together imply that

$$I = \int_1^\infty f(x)\,\mathrm{d}x \geq \int_1^\infty g(x)\,\mathrm{d}x = \infty,$$

which is what we are trying to show.

To complete the reasoning we have to find the easy-to-integrate function $g(x)$. Based on what we have done above our first guess would be $g(x) = \frac{1}{x}$, but this does not work, since

$$\frac{x}{x^2+1} < \frac{x}{x^2} = \frac{1}{x}.$$

So with this choice of $g(x)$ we get $g(x) > f(x)$ instead of $g(x) < f(x)$.

One way to simplify $f(x)$ and get a smaller function is to remember that *by increasing the denominator in a fraction we decrease the fraction.* Thus, for $x > 1$ we have

$$f(x) = \frac{x}{x^2+1} > \frac{x}{x^2+x^2} = \frac{x}{2x^2} = \frac{1}{2x}.$$

So we let $g(x) = \frac{1}{2x}$. Then we find

$$\begin{aligned} I &= \int_1^\infty \frac{x}{x^2+1}\,\mathrm{d}x \\ &\geq \int_1^\infty \frac{1}{2x}\,\mathrm{d}x \\ &= \infty. \end{aligned}$$

5.3. The Tail Theorem. *If $y = f(x)$ is a continuous function for $x \geq a$, and if $b > a$ then*

$$\int_a^\infty f(x)\,\mathrm{d}x \textit{ exists if and only if } \int_b^\infty f(x)\,\mathrm{d}x \textit{ exists}$$

Moreover, if these integrals exist, then (18) holds: $\int_a^\infty f(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x + \int_b^\infty f(x)\,\mathrm{d}x$.

PROOF. For any finite M one has

$$\int_a^M f(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x + \int_b^M f(x)\,\mathrm{d}x$$

The theorem follows by taking the limit for $M \to \infty$ on both sides of this equation. $\square$

The following two examples show how one can use this fact.

5.3.1. *Example. Does the integral*

$$I = \int_0^\infty \frac{\mathrm{d}x}{x^3+1}$$

exist?

The integrand here is the same as in § 5.2.1 where we found that

$$\int_1^\infty \frac{\mathrm{d}x}{x^3+1} < \frac{1}{2},$$

and in particular is finite. Since the function $f(x) = 1/(x^3+1)$ is continuous for $0 \le x \le 1$ we may conclude that

$$I = \int_0^\infty \frac{\mathrm{d}x}{x^3+1} = \int_0^1 \frac{\mathrm{d}x}{x^3+1} + \int_1^\infty \frac{\mathrm{d}x}{x^3+1}$$

so that I is finite.

If we want an upper estimate for the integral we can use the fact that we already know that the integral from 1 to ∞ is not more than $\frac{1}{2}$ and estimate the integral from $x = 0$ to $x = 1$. For $0 \le x \le 1$ we have

$$x^3 + 1 \ge 1 \implies \frac{1}{x^3+1} \le 1$$

This implies

$$\int_0^1 \frac{\mathrm{d}x}{x^3+1} \le \int_0^1 1\,\mathrm{d}x = 1,$$

and hence

$$\int_0^\infty \frac{\mathrm{d}x}{x^3+1} = \int_0^1 \frac{\mathrm{d}x}{x^3+1} + \int_1^\infty \frac{\mathrm{d}x}{x^3+1} \le 1 + \frac{1}{2} = \frac{3}{2}.$$

5.3.2. *The area under the bell curve.* The ***bell curve*** which plays a central role in probability and statistics, is the graph of the function

$$n(x) = ae^{-x^2/b},$$

where a and b are positive constants whose values will not be important in this example. In fact, to simplify this example we will choose them to be $a = 1$ and $b = 1$ so that we are dealing with the function $n(x) = e^{-x^2}$. The question now is, *is the area under the bell curve finite, and can we estimate it?* In terms of improper integrals, we want to know if the integral

$$A = \int_{-\infty}^\infty e^{-x^2}\,\mathrm{d}x$$

exists.

We write the integral as the sum of two integrals

$$\int_{-\infty}^\infty e^{-x^2}\,\mathrm{d}x = \int_{-\infty}^0 e^{-x^2}\,\mathrm{d}x + \int_0^\infty e^{-x^2}\,\mathrm{d}x$$

Since the function $n(x) = e^{-x^2}$ is even these two integrals are equal, so if we can show

$n(x) = n(-x)$ so the bell curve is symmetric that one of them exists, the other will also exist.

When x is large x^2 is much larger than x so that e^{-x^2} will be smaller than e^{-x}; since e^{-x} is a function we know how to integrate it seems like a good idea to compare the bell curve with the graph of e^{-x}. To make the comparison (and to use the Comparison Theorem) we first check if $e^{-x^2} \le e^{-x}$ really is true. We find that this is true if $x \ge 1$, namely:

$$x \ge 1 \implies x^2 \ge x \implies -x^2 \le -x \implies e^{-x^2} \le e^{-x}.$$

We can therefore use e^{-x} to estimate the integral of e^{-x^2} for $x \ge 1$:

$$\int_1^\infty e^{-x^2}\,\mathrm{d}x \le \int_1^\infty e^{-x}\,\mathrm{d}x = \lim_{M\to\infty}\Big[-e^{-x}\Big]_1^M = \lim_{M\to\infty}\left(-e^{-M}+e^{-1}\right) = e^{-1}.$$

Between $x=0$ and $x=1$ we have $e^{-x^2} \le 1$, so

$$\int_0^1 e^{-x^2}\,\mathrm{d}x \le \int_0^1 1\,\mathrm{d}x = 1,$$

and therefore

$$\int_0^\infty e^{-x^2}\,\mathrm{d}x = \int_0^1 e^{-x^2}\,\mathrm{d}x + \int_1^\infty e^{-x^2}\,\mathrm{d}x \le 1 + \frac{1}{e}.$$

Since the bell curve is symmetric, we get the same estimates for the integral over $-\infty < x \le 0$:

$$\int_{-\infty}^0 e^{-x^2}\,\mathrm{d}x \le 1 + \frac{1}{e}.$$

In the end we have shown that the area under the bell curve is finite, and that it is bounded by

$$A = \int_{-\infty}^\infty e^{-x^2}\,\mathrm{d}x \le 2 + \frac{2}{e}.$$

With quite a bit more work, and using an indirect argument one can actually compute the integral (without finding an antiderivative of e^{-x^2}). The true value turns out to be

$$A = \int_{-\infty}^\infty e^{-x^2}\,\mathrm{d}x = \sqrt{\pi}.$$

For comparison,

$$\sqrt{\pi} \approx 1.772\,453\,850\,9\ldots \text{ and } 2 + \frac{2}{e} \approx 2.735\,758\,882\,3\ldots.$$

6. Problems

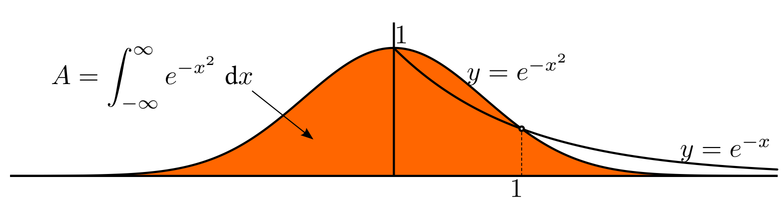


Figure 5. Estimating the area under the bell curve.

Which of the following inequalities are true for all $x > a$ (you get to choose a):

1. $\dfrac{1}{x} > \dfrac{1}{x^2}$?
Answer: True. If $x > 1$ then $x^2 > x$ and therefore $\frac{1}{x^2} < \frac{1}{x}$. So the inequality is true if $x > a$ where we choose $a = 1$.

2. $\dfrac{1}{x-1} < \dfrac{1}{x}$?

3. $\dfrac{x}{x^2+2x} > \dfrac{1}{x}$?

4. $\dfrac{x}{\sqrt{x^3-1}} > x^{-1/2}$?

5. $\dfrac{2x^2-3x+1}{(x^2-x)^3} < 2x^{-4}$?

Which of the following inequalities are true for all x with $0 < x < a$ (you get to choose a again, as long as $a > 0$):

6. $\dfrac{3}{x} > \dfrac{3}{x^2}$?
Answer: False. If $0 < x < 1$ then $x^2 < x$ and therefore $\frac{1}{x^2} > \frac{1}{x}$, which implies $\frac{3}{x^2} > \frac{3}{x}$. So, more precisely, the inequality is false if $0 < x < a$ where we choose $a = 1$.

7. $\dfrac{x}{x^2+x} < \dfrac{1}{x}$?

8. $\dfrac{2}{x-x^2} < \dfrac{2}{x}$?

9. $\dfrac{2}{x+x^2} < \dfrac{2}{x}$?

10. $\dfrac{4}{x+20x^2} < \dfrac{2}{x}$?

11. $\dfrac{2}{x-x^2} < \dfrac{4}{x}$?

For each of the following improper integrals draw the graph of the integrand, and decide if the integral exists. If it does try to find an estimate for how large the integral is.

12. $\displaystyle\int_0^\infty \frac{u^2}{\left(u^2+1\right)^2}\,\mathrm{d}u$

13. $\displaystyle\int_0^\infty \frac{u^3}{\left(u^2+1\right)^2}\,\mathrm{d}u$

14. $\displaystyle\int_0^\infty \frac{u^4}{\left(u^2+1\right)^2}\,\mathrm{d}u$

15. $\displaystyle\int_\pi^\infty \frac{\sin x}{x^2}\,\mathrm{d}x$

16. $\displaystyle\int_0^\infty \frac{\sin(x^2)}{x^2}\,\mathrm{d}x$ (what happens at $x = 0$?)

17. **[Group Problem]** (*Fresnel integrals from optics.*) For background on Fresnel integrals read the Wikipedia article on the subject.

Consider the function $F(x) = \dfrac{\sin x^2}{2x}$.

(a) Compute $\lim_{x\to 0} F(x)$ and $\lim_{x\to\infty} F(x)$.

(b) Show that $F'(x) = \cos x^2 - \dfrac{\sin x^2}{2x^2}$.

(c) Show that the integral $\displaystyle\int_0^\infty \frac{\sin x^2}{2x^2}\,\mathrm{d}x$ exists.

(d) Show that

$$\int_0^\infty \cos x^2\,\mathrm{d}x = \int_0^\infty \frac{\sin x^2}{2x^2}\,\mathrm{d}x.$$

(e) True or False: *if for some function f the improper integral $\int_0^\infty f(x)\,\mathrm{d}x$ exists then it must be true that* $\lim_{x\to\infty} f(x) = 0$*?*

18. **[Group Problem]** Suppose

$$p(x) = e^{-ax},$$

where $a > 0$ is a constant.

(a) Show that $\int_0^\infty p(x)\,\mathrm{d}x < \infty$. (hint: you can compute the integral.)

(b) Compute

$$\frac{\int_0^\infty xp(x)\,\mathrm{d}x}{\int_0^\infty p(x)\,\mathrm{d}x}$$

by integrating by parts in the integral in the numerator.

(c) Compute

$$\frac{\int_0^\infty x^2p(x)\,\mathrm{d}x}{\int_0^\infty p(x)\,\mathrm{d}x}$$

by integrating by parts twice in the integral in the numerator.

19. **[Group Problem]** Suppose

$$p(x) = e^{-x^2/\sigma^2}$$

where $\sigma > 0$ is a constant.
(σ is the Greek lower case "s," pronounced as "sigma.")

(a) Show that $\int_{-\infty}^{\infty} p(x)\,\mathrm{d}x < \infty$. (hint: there is no antiderivative of e^{-x^2/σ^2}, so don't look for it. Instead follow example 5.3.2.)

(b) Compute

$$\frac{\int_{-\infty}^{\infty} xp(x)\,\mathrm{d}x}{\int_0^{\infty} p(x)\,\mathrm{d}x}.$$

Hint: there is an easy antiderivative for the integral in the numerator. Do that one first.

(c) Compute

$$\frac{\int_{-\infty}^{\infty} x^2p(x)\,\mathrm{d}x}{\int_0^{\infty} p(x)\,\mathrm{d}x}$$

by integrating by parts in the integral in the numerator.

20. [Group Problem]

"The area under the bell curve is $\frac{1}{2}$-factorial"

In this problem we look at a new function $F(n)$. For any number n we would like to define

$$F(n) = \int_0^{\infty} t^n e^{-t}\,\mathrm{d}t.$$

The integral that defines $F(n)$ is improper so it is not automatically clear that $F(n)$ exists.

(a) Show that the integral exists if $n \geq 0$.

(b) Compute $F(0)$, $F(1)$, and $F(2)$.

(c) Use integration by parts to show that $F(n+1) = (n+1)F(n)$ Holds for any $n \geq 0$. Use this to Compute $F(10)$.

(d) If you set $n = -1$ in the relation $F(n+1) = (n+1)F(n)$ then you get $F(0) = 0 \cdot F(-1) = 0$. This should contradict the value for $F(0)$ you found in **(b)**. What is wrong?

(e) Show that the integral for $F(\frac{1}{2})$ exists, and use a substitution to show that

$$F(\tfrac{1}{2}) = \int_{-\infty}^{\infty} e^{-u^2}\,\mathrm{d}u.$$

How does this justify the statement at the beginning of the problem?

CHAPTER 3

First order differential Equations

1. What is a Differential Equation?

A ***differential equation*** is an equation involving an unknown function and its derivatives. A general differential equation can contain second derivatives and higher derivatives of the unknown function, but in this course we will only consider differential equations that contain first order derivatives. So the most general differential equation we will look at is of the form

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x, y). \tag{20}$$

Here $f(x, y)$ is any expression that involves both quantities x and y. For instance, if $f(x, y) = x^2 + y^2$ then the differential equation represented by (20) is

$$\frac{\mathrm{d}y}{\mathrm{d}x} = x^2 + y^2.$$

In this equation x is a variable, while y is a function of x.

Differential equations appear in many science and engineering problems, as we will see in the section on applications. But first let's think of a differential equation as a purely mathematical question: *which functions $y = y(x)$ satisfy the equation* (20)*?* It turns out that there is no general method that will always give us the answer to this question, but there are methods that work for certain special kinds of differential equations. To explain all this, this chapter is divided into the following parts:

- Some basic examples that give us a clue as to what the solution to a general differential equation will look like.
- Two special kinds of differential equation ("separable" and "linear"), and how to solve them.
- How to visualize the solution of a differential equation (using "direction fields") and how to compute the solution with a computer using Euler's method.
- Applications: a number of examples of how differential equations come up and what their solutions mean.

2. Two basic examples

2.1. Equations where the RHS does not contain y. Which functions $y = y(x)$ satisfy

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \sin x\ ? \tag{21}$$

This is a differential equation of the form (20) where the function f that describes the Right Hand Side is given by $f(x, y) = \sin x$. In this example the function f does not depend on the unknown function y. Because of this the differential equation really asks "*which functions of x have* $\sin x$ *as derivative?*" In other words, which functions are the antiderivative of $\sin x$? We know the answer, namely

$$y = \int \sin x \, \mathrm{d}x = -\cos x + C$$

where C is an arbitrary constant. This is the solution to the differential equation (21). This example shows us that there is not just one solution, but that there are many solutions. The expression that describes ***all solutions*** to the differential equation (21) is called ***the general solution***. It contains an unknown constant C that is allowed to have arbitrary values.

To give meaning to the constant C we can observe that when $x = 0$ we have

$$y(0) = -\cos 0 + C = -1 + C.$$

So the constant C is nothing but

$$C = y(0) + 1.$$

For instance, the solution of (21) that also satisfies $y(0) = 4$ has $C = 4 + 1 = 5$, and thus is given by

$$y(x) = -\cos x + 5.$$

We have found that there are many solutions to the differential equation (21) (because of the undetermined constant C), but as soon as we prescribe the value of the solution for one value of x, such as $x = 0$, then there is exactly one solution (because we can compute the constant C.)

2.2. The exponential growth example. *Which functions equal their own derivative*, i.e. which functions satisfy

$$\frac{\mathrm{d}y}{\mathrm{d}x} = y \, ?$$

Everyone knows at least one example, namely $y = e^x$. But there are more solutions: the function $y = 0$ also is its own derivative. From the section on exponential growth in math 221 we know all solutions to $\frac{\mathrm{d}y}{\mathrm{d}x} = y$. They are given by

$$y(x) = Ce^x,$$

where C can be an arbitrary number. If we know the solution $y(x)$ for some value of x, such as $x = 0$, then we can find C by setting $x = 0$:

$$y(0) = C.$$

Again we see that instead of there being one solution, the general solution contains an arbitrary constant C.

2.3. Summary. The two examples that we have just seen show us that for certain differential equations

- there are many solutions,
- the formula for the general solution contains an undetermined constant C,
- the undetermined constant C becomes determined once we specify the value of the solution y at one particular value of x.

It turns out that these features are found in almost all differential equations of the form (20). In the next two sections we will see methods for computing the general solution to two frequently occurring kinds of differential equation, the *separable equations*, and the *linear equations.*

3. First Order Separable Equations

By definition a ***separable differential equation*** is a diffeq of the form

$$y'(x) = F(x)G(y(x)), \quad \text{or} \quad \frac{\mathrm{d}y}{\mathrm{d}x} = F(x)G(y). \tag{22}$$

Thus the function $f(x, y)$ on the right hand side in (20) has the special form

$$f(x, y) = F(x)\, G(y).$$

For example, the differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \sin(x)\big(1 + y^2\big)$$

is separable, and one has $F(x) = \sin x$ and $G(y) = 1 + y^2$. On the other hand, the differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} = x + y$$

is not separable.

3.1. Solution method for separable equations.

To solve this equation divide by $G(y(x))$ to get

$$\frac{1}{G(y(x))}\frac{\mathrm{d}y}{\mathrm{d}x} = F(x). \tag{23}$$

Next find a function $H(y)$ whose derivative with respect to y is

$$H'(y) = \frac{1}{G(y)} \qquad \left(\text{solution: } H(y) = \int \frac{dy}{G(y)}.\right) \tag{24}$$

Then the chain rule implies that the left hand side in (23) can be written as

$$\frac{1}{G(y(x))}\frac{\mathrm{d}y}{\mathrm{d}x} = H'(y(x))\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}H(y(x))}{\mathrm{d}x}.$$

Thus (23) is equivalent with

$$\frac{\mathrm{d}H(y(x))}{\mathrm{d}x} = F(x).$$

In words: $H(y(x))$ is an antiderivative of $F(x)$, which means we can find $H(y(x))$ by integrating $F(x)$:

$$H(y(x)) = \int F(x)dx + C. \tag{25}$$

Once we have found the integral of $F(x)$ this gives us $y(x)$ in implicit form: the equation (25) gives us $y(x)$ as an *implicit function* of x. To get $y(x)$ itself we must solve the equation (25) for $y(x)$.

A quick way of organizing the calculation goes like this:

> To solve $\dfrac{\mathrm{d}y}{\mathrm{d}x} = F(x)G(y)$ we first *separate the variables,*
>
> $$\frac{\mathrm{d}y}{G(y)} = F(x)\,\mathrm{d}x,$$
>
> and then integrate,
>
> $$\int \frac{\mathrm{d}y}{G(y)} = \int F(x)\,\mathrm{d}x.$$

The result is an implicit equation for the solution y with one undetermined integration constant.

Determining the constant. The solution we get from the above procedure contains an arbitrary constant C. If the value of the solution is specified at some given x_0, i.e. if $y(x_0)$ is known then we can express C in terms of $y(x_0)$ by using (25).

3.2. A snag: We have to divide by $G(y)$ which is problematic when $G(y) = 0$. This has as consequence that in addition to the solutions we found with the above procedure, there are at least a few more solutions: the zeroes of $G(y)$ (see Example 3.4 below). In addition to the zeroes of $G(y)$ there sometimes can be more solutions, as we will see in Example 10.2 on "Leaky Bucket Dating."

3.3. Example. We solve

$$\frac{\mathrm{d}z}{\mathrm{d}t} = (1 + z^2)\cos t.$$

Separate variables and integrate

$$\int \frac{\mathrm{d}z}{1+z^2} = \int \cos t\,\mathrm{d}t,$$

to get

$$\arctan z = \sin t + C.$$

Finally solve for z and we find the general solution

$$z(t) = \tan\bigl(\sin(t) + C\bigr).$$

3.4. Example: the snag in action. If we apply the method to $y'(x) = y$, we get

$$y(x) = e^{x+C}.$$

No matter how we choose C we never get the function $y(x) = 0$, even though $y(x) = 0$ satisfies the equation. This is because here $G(y) = y$, and $G(y)$ vanishes for $y = 0$.

4. Problems

For each of the following differential equations

- find the general solution,
- indicate which, if any, solutions were lost while separating variables,
- find the solution that satisfies the indicated initial values.

1. $\dfrac{\mathrm{d}y}{\mathrm{d}x} = xy, \quad y(2) = -1.$

2. $\dfrac{\mathrm{d}y}{\mathrm{d}x} + x\cos^2 y = 0, \quad y(0) = \frac{\pi}{3}.$

3. $\dfrac{\mathrm{d}y}{\mathrm{d}x} + \dfrac{1+x}{1+y} = 0, \quad y(0) = A.$

4. $y^2\dfrac{\mathrm{d}y}{\mathrm{d}x} + x^3 = 0, y(0) = A.$

5. $\dfrac{\mathrm{d}y}{\mathrm{d}x} + 1 - y^2 = 0, \quad y(0) = A.$

6. $\dfrac{\mathrm{d}y}{\mathrm{d}x} + 1 + y^2 = 0, \quad y(0) = A.$

7. $\dfrac{\mathrm{d}y}{\mathrm{d}x} + \dfrac{x^2-1}{y} = 0, \quad y(0) = 1.$

8. **[Group Problem]** Let $P(t)$ be the size of a colony of bacteria in a Petri dish in some experiment. Assume that the size of the colony changes according to the so-called logistic equation:

$$\frac{\mathrm{d}P}{\mathrm{d}t} = \frac{1}{50}P(1000 - P),$$

Assume also that in the beginning of the experiment the population size is $P = 100$.

(a) Find the general solution to the differential equation.

(b) Find the solution that satisfies the given initial conditions.

(c) How long does it take the population to reach size $P = 500$?

(d) How long does it take the population to reach size $P = 900$?

(e) What value does P have when $\frac{\mathrm{d}P}{\mathrm{d}t}$ is the largest (hint: you do not need to solve the differential equation–this question has a very short answer.)

5. First Order Linear Equations

Differential equations of the form equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} + a(x)y = k(x) \tag{26}$$

are called ***first order linear***.

5.1. The Integrating Factor.

Linear equations can always be solved by multiplying both sides of the equation with a specially chosen function called the ***integrating factor***. It is defined by

$$A(x) = \int a(x)\,\mathrm{d}x, \qquad m(x) = e^{A(x)}. \tag{27}$$

Here $m(x)$ is the integrating factor. It looks like we just pulled this definition of $A(x)$ and $m(x)$ out of a hat. The example in § 5.2 shows another way of finding the integrating factor, but for now let's go on with these two functions.

Multiply the equation (26) by the integrating factor $m(x)$ to get

$$m(x)\frac{\mathrm{d}y}{\mathrm{d}x} + a(x)m(x)y = m(x)k(x).$$

By the chain rule the integrating factor satisfies

$$\frac{\mathrm{d}m(x)}{\mathrm{d}x} = \frac{\mathrm{d}\,e^{A(x)}}{\mathrm{d}x} = \underbrace{A'(x)}_{=a(x)}\ \underbrace{e^{A(x)}}_{=m(x)} = a(x)m(x).$$

Therefore one has

$$\begin{aligned}\frac{\mathrm{d}m(x)y}{\mathrm{d}x} &= m(x)\frac{\mathrm{d}y}{\mathrm{d}x} + a(x)m(x)y \\ &= m(x)\Big\{\frac{\mathrm{d}y}{\mathrm{d}x} + a(x)y\Big\} \\ &= m(x)k(x).\end{aligned}$$

Integrating and then dividing by the integrating factor gives the solution

$$y = \frac{1}{m(x)}\left(\int m(x)k(x)\,\mathrm{d}x + C\right).$$

In this derivation we have to divide by $m(x)$, but since $m(x) = e^{A(x)}$ and since exponentials never vanish we know that $m(x) \neq 0$, so we can always divide by $m(x)$.

5.2. An example. *Find the general solution to the differential equation*

$$\frac{\mathrm{d}y}{\mathrm{d}x} = y + x.$$

Then find the solution that satisfies

$$y(2) = 0. \tag{28}$$

Solution. We first write the equation in the standard linear form

$$\frac{\mathrm{d}y}{\mathrm{d}x} - y = x, \tag{29}$$

and then multiply with the integrating factor $m(x)$. We could of course memorize the formulas (27) that lead to the integrating factor, but a safer approach is to remember the following procedure, which will always give us the integrating factor.

Assuming that $m(x)$ is as yet unknown we multiply the differential equation (29) with m,

$$m(x)\frac{\mathrm{d}y}{\mathrm{d}x} - m(x)y = m(x)x. \tag{30}$$

If $m(x)$ is such that

$$-m(x) = \frac{\mathrm{d}m(x)}{\mathrm{d}x}, \tag{31}$$

then equation (30) implies

$$m(x)\frac{\mathrm{d}y}{\mathrm{d}x} + \frac{\mathrm{d}m(x)}{\mathrm{d}x}y = m(x)x.$$

The expression on the left is exactly what comes out of the product rule – this is the point of multiplying with $m(x)$ and then insisting on (31). So, if $m(x)$ satisfies (31), then the differential equation for y is equivalent with

$$\frac{\mathrm{d}m(x)y}{\mathrm{d}x} = m(x)x.$$

We can integrate this equation,

$$m(x)y = \int m(x)x \, \mathrm{d}x,$$

and thus find the solution

$$y(x) = \frac{1}{m(x)} \int m(x)x \, \mathrm{d}x. \tag{32}$$

All we have to do is find the integrating factor m. This factor can be any function that satisfies (31). Equation (31) is a differential equation for m, but it is separable, and we can easily solve it:

$$\frac{\mathrm{d}m}{\mathrm{d}x} = -m \iff \frac{\mathrm{d}m}{m} = -\mathrm{d}x \iff \ln|m| = -x + C.$$

Since *we only need one integrating factor* m we are not interested in finding all solutions of (31), and therefore we can choose the constant C. The simplest choice is $C = 0$, which leads to

$$\ln|m| = -x \iff |m| = e^{-x} \iff m = \pm e^{-x}.$$

Again, we only need one integrating factor, so we may choose the $\pm$ sign: the simplest choice for m here is

$$m(x) = e^{-x}.$$

With this choice of integrating factor we can now complete the calculation that led to (32). The solution to the differential equation is

$$\begin{aligned} y(x) &= \frac{1}{m(x)} \int m(x)x \, \mathrm{d}x \\ &= \frac{1}{e^{-x}} \int e^{-x} x \, \mathrm{d}x && \text{(integrate by parts)} \\ &= e^{x} \left\{ -e^{-x}x - e^{-x} + C \right\} \\ &= -x - 1 + Ce^{x}. \end{aligned}$$

This is the general solution.

To find the solution that satisfies not just the differential equation, but also the "initial condition" (28), i.e. $y(2) = 0$, we compute $y(2)$ for the general solution,

$$y(2) = -2 - 1 + Ce^2 = -3 + Ce^2.$$

The requirement $y(2) = 0$ then tells us that $C = 3e^{-2}$. The solution of the differential equation that satisfies the prescribed initial condition is therefore

$$y(x) = -x - 1 + 3e^{x-2}.$$

6. Problems

1. In example 5.2 we needed a function $m(x)$ that satisfies (31). The function $m(x) = 0$ satisfies that equation. Why did we not choose $m(x) = 0$?

2. Why can't we simplify the computation at the end in example 5.2 by canceling the two factors $m(x)$ as follows:

$$\begin{aligned} y(x) &= \frac{1}{\cancel{m(x)}} \int \cancel{m(x)}\, x \, \mathrm{d}x \\ &= \int x \, \mathrm{d}x \\ &= \frac{1}{2}x^2 + C? \end{aligned}$$

For each of the following differential equations

- *specify the differential equation that the integrating factor satisfies,*
- *find one integrating factor,*
- *find the general solution,*
- *find the solution that satisfies the specified initial conditions.*

In these problems K and N are constants.

3. $\dfrac{\mathrm{d}y}{\mathrm{d}x} = -y + x, \quad y(0) = 0.$

4. $\dfrac{\mathrm{d}y}{\mathrm{d}x} = 2y + x^2, \quad y(0) = 0.$

5. $\dfrac{\mathrm{d}y}{\mathrm{d}x} + 2y + e^{x} = 0.$ •

6. $\dfrac{\mathrm{d}y}{\mathrm{d}x} - (\cos x)y = e^{\sin x}, y(0) = A.$ •

7. $\dfrac{\mathrm{d}y}{\mathrm{d}x} = -10y + e^{-x}, \quad y(0) = 0.$

8. $\dfrac{\mathrm{d}y}{\mathrm{d}x} = y \tan x + 1, \quad y(0) = 0.$ •

9. $\dfrac{\mathrm{d}y}{\mathrm{d}x} = -y \tan x + 1, \quad y(0) = 0.$ •

10. $\cos^2 x \dfrac{\mathrm{d}y}{\mathrm{d}x} = N - y \quad y(0) = 0.$ •

11. $x\dfrac{\mathrm{d}y}{\mathrm{d}x} = y + x, \quad y(2) = 0.$ •

12. $\dfrac{\mathrm{d}y}{\mathrm{d}x} = -xy + x^3, \quad y(0) = 0.$

13. $\dfrac{\mathrm{d}y}{\mathrm{d}x} = -y + \sin x, \quad y(0) = 0.$

14. $\dfrac{\mathrm{d}y}{\mathrm{d}x} = -Ky + \sin x, \quad y(0) = 0.$

15. $\dfrac{\mathrm{d}y}{\mathrm{d}x} + x^2 y = 0, \quad y(1) = 5.$ •

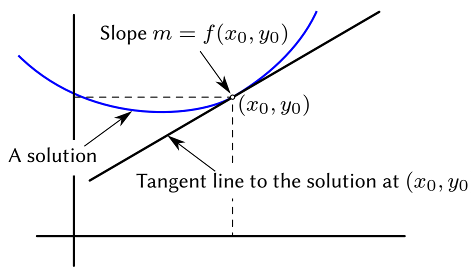
16. $\dfrac{\mathrm{d}y}{\mathrm{d}x} + (1 + 3x^2)y = 0, \quad y(1) = 1.$ •

7. Direction Fields

We can visualize a differential equation by drawing the corresponding direction field. Consider a differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} = f(x, y)$$

where f is some given function. The differential equation tells us that if we know a point (x_0, y_0) on the graph of a solution $y = y(x)$ of the differential equation, then we also know the slope of the graph at that point. We can draw the tangent line to the graph of the solution:



If we have not yet solved the differential equation then we don't know any points on the solution. In this case we can sample the xy-plane, compute $f(x, y)$ at all our sample points, and draw the tangent lines a solution would have if it passed through one of the sample points. In Figure 1 this is done for the differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -y + \sin \pi x.$$

The direction field on the left in Figure 1 gives us an idea of what the solutions should look like. Whenever a solution passes through one of the sample points the slope of its graph is given by the line segment drawn at that sample point. It is clear from a quick look at Figure 1 that drawing a direction field involves computing $f(x, y)$ (i.e. $-y + \sin \pi x$ in our example) for *many, many* points (x, y). This kind of repetitive computation is better

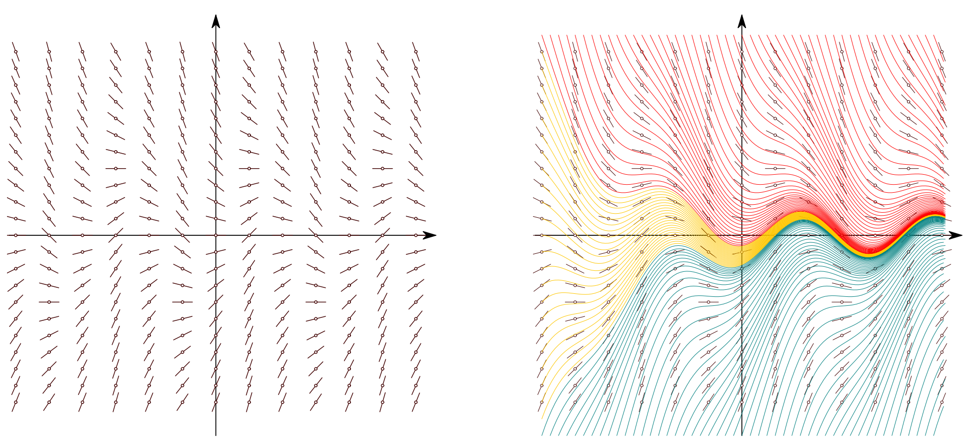


Figure 1. Direction field for $\frac{\mathrm{d}y}{\mathrm{d}x} = -y + \sin \pi x$ on the left, and the same direction field with solutions to the differential equation.

done by a computer, and an internet search quickly leads to a number of websites that produce direction fields for our favorite differential equation. In particular the ODE page at the Virtual Math Museum of UC Irvine,

http://virtualmathmuseum.org

draws both direction fields and approximations to solutions found using Euler's method, to which we now turn.

8. Euler's method

8.1. The idea behind the method. Consider again the differential equation (20),

$$\frac{dy}{dx} = f(x, y).$$

Suppose we know one point on a solution, i.e. suppose we know that a given solution to this equation satisfies $y = y_0$ when $x = x_0$, i.e.

$$y(x_0) = y_0. \tag{33}$$

The differential equation then tells us what the derivative of the solution is at $x = x_0$, namely,

$$y'(x_0) = f(x_0, y_0).$$

The definition a derivative says that

$$y'(x_0) = \lim_{h\to 0} \frac{y(x_0 + h) - y(x_0)}{h}$$

so that we have

$$\lim_{h\to 0} \frac{y(x_0 + h) - y(x_0)}{h} = f(x_0, y_0). \tag{34}$$

Keep in mind that the right hand side is what we get by substituting the x and y values that we know for the solution in f. So if we know x_0 and y_0 then we can also compute $f(x_0, y_0)$.

If we don't know the solution then we cannot compute the left hand side in (34), but, following Euler, we can make an approximation. If instead of letting $h \to 0$ we choose a small number $h > 0$, then we may assume that

$$\frac{y(x_0 + h) - y(x_0)}{h} \approx f(x_0, y_0). \tag{35}$$

Here "$\approx$" means "approximately equal," which is a vaguely defined concept. It means that the difference between the two quantities in (35) is "small" and we will not worry too much about the error in (35) (such issues are addressed in more advanced courses on Numerical Analysis; e.g. Math 514 at UW Madison). In the approximate equation (35) all quantities are known except $y(x_0 + h)$. After solving (35) for $y(x_0 + h)$ we find

$$y(x_0 + h) \approx y_0 + hf(x_0, y_0). \tag{36}$$

Euler's idea (see Figure 2) was to forget that this is an approximation, and declare that we now know a new point on the graph of the solution, namely

$$x_1 = x_0 + h, \qquad y_1 = y_0 + hf(x_0, y_0). \tag{37}$$

Assuming that our solution satisfies $y(x_1) = y_1$ exactly (rather than just approximately), we can apply the same procedure and get another new point on the graph of our solution:

$$x_2 = x_1 + h, \qquad y_2 = y_1 + hf(x_1, y_1).$$

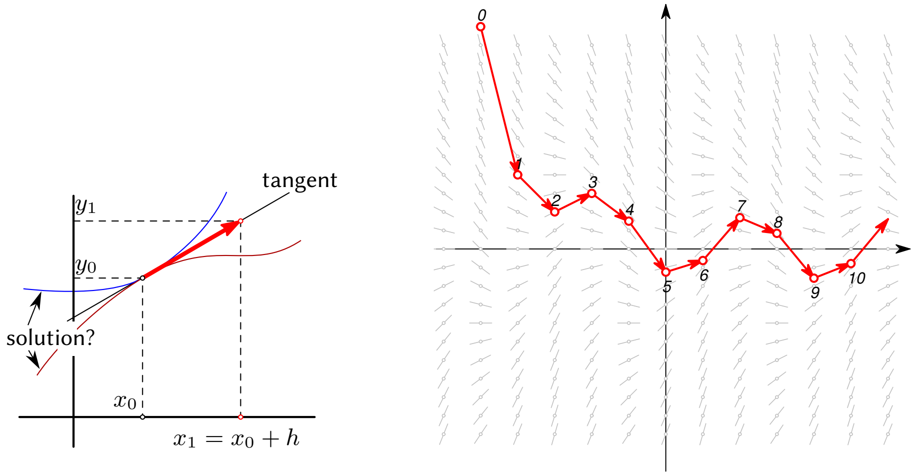


Figure 2. Approximating a solution with Euler's method. *On the left*: one step of Euler's method. Knowing one point (x_0, y_0) on the graph does not immediately tell us what the solution is, but it does tell us what the tangent to the graph at (x_0, y_0) is. We can then guess that the point on the tangent with $x_1 = x_0 + h$ and $y_1 = y_0 + f(x_0, y_0)h$ *almost* lies on the graph.
On the right: repeating the procedure gives us a sequence of points that approximate the solution. Reducing the "step size" h should give better approximations.

By repeating this procedure we can generate a whole list of points (x_0, y_0), (x_1, y_1), (x_2, y_2), (x_3, y_3), etc…that lie approximately on the graph of the solution passing through the given initial point (x_0, y_0).

8.2. Setting up the computation. If we are given x_{start} and $y(x_{\text{start}}) = y_{\text{start}}$, and if we want to find $y(x_{\text{end}})$ for some $x_{\text{end}} > x_{\text{start}}$, then we can decide to approximate $y(x_{\text{end}})$ by applying n steps of Euler's method. Since each step advances x by an amount h, application of Euler's method will advance x by nh after n steps. Thus we want

$$nh = x_{\text{end}} - x_{\text{start}}, \quad \text{i.e.} \quad h = \frac{x_{\text{end}} - x_{\text{start}}}{n}.$$

Starting with the known value of y at x_{start} we then apply Euler's method n times by filling out a table of the following form:

x_k	y_k	$m_k = f(x_k, y_k)$	$y_{k+1} = y_k + m_k \cdot h$
$x_0 = x_{\text{start}}$	y_{start}	m_0	y_1
$x_1 = x_0 + h$	y_1	m_1	y_2
$x_2 = x_0 + 2h$	y_2	m_2	y_3
$x_3 = x_0 + 3h$	y_3	m_3	y_4
$\vdots$			$\vdots$
$x_{n-1} = x_0 + (n-1)h$	y_{n-1}	m_{n-1}	y_n
$x_n = x_{\text{end}}$	y_n		

The procedure is very mechanical and repetitive, and is best programmed on a computer.

Once the table has been computed the values in the first two columns can be used to graph the approximation to the real solution.

9. Problems

1. Let $y(t)$ be the solution of

$$\frac{\mathrm{d}y}{\mathrm{d}t} = -t \cdot y, \quad y(1.0) = 10.0.$$

We want to compute $y(3.0)$.

(a) Find an exact formula for $y(3.0)$ by solving the equation (the equation is both separable and linear, so we have at least two methods for finding the solution.)

(b) If we use Euler's method with step size $h = 2.0$, then how many steps do we have to take to approximate $y(3.0)$? Compute the approximation with $h = 2.0$.

(c) Find the approximations if $h = 1.0$, $h = 2/3$, and $h = 0.5$. Organize the computations following the table in 8.2.

(d) Compare the results from the computations in **(b)** and **(c)** with the true value of $y(3.0)$ from part **(a)**.

2. The function $y(x) = e^x$ is the solution to

$$\frac{\mathrm{d}y}{\mathrm{d}x} = y, \quad y(0) = 1.$$

(a) Approximate $e = y(1)$ by using Euler's method first with step size $h = 1$, then with $h = 1/2$, and then with $h = 1/3$. What are the approximations you find for $y(1)$ in each case?

(b) Look for a pattern in the answers to **(a)**.

(c) Find a formula for the result of applying Euler's method n times with step size $h = \frac{1}{n}$.

3. Use Euler's method to approximate the solution to

$$\frac{\mathrm{d}y}{\mathrm{d}x} = -10y, \quad y(0) = 1,$$

and, in particular to find $y(1)$. Use various step sizes. How small do you have to make the step size before the answer seems reasonable? Explain.

10. Applications of Differential Equations

Differential equations are very often used to describe how some "object" or "system" changes or evolves in time. If the object or system is simple enough then its state is completely determined by one number (say y) that changes with time t.

A differential equation for the system tells us how the system changes in time, by specifying the rate of change of the state $y(t)$ of the system. This rate of change can depend on time, and it can depend on the state of the system, i.e. on $y(t)$. This dependence can be expressed as an equation of the form

$$\frac{\mathrm{d}y}{\mathrm{d}t} = f(y, t). \tag{38}$$

The function f describes the *evolutionary law* of our system (synonyms: "evolutionary law", "dynamical law", "evolution equation for y").

10.1. Example: carbon dating.

Suppose we have a fossil, and we want to know how old it is.

All living things contain carbon, which naturally occurs in two isotopes, C_{14} (unstable) and C_{12} (stable). A long as the living thing is alive it eats & breaths, and its ratio of C_{12} to C_{14} is kept constant. Once the thing dies the isotope C_{14} decays into C_{12} at a steady rate that is proportional to the amount of C_{14} it contains.

Let $y(t)$ be the ratio of C_{14} to C_{12} at time t. The law of radioactive decay says that there is a constant $k > 0$ such that

$$\frac{\mathrm{d}y(t)}{\mathrm{d}t} = -ky(t).$$

Solve this differential equation (it is both separable and first order linear: either method works) to find the general solution

$$y(t; C) = Ce^{-kt}.$$

After some lab work it is found that the current C_{14}/C_{12} ratio of our fossil is y_{now}. Thus we have

$$y_{\text{now}} = Ce^{-kt_{\text{now}}} \implies C = y_{\text{now}}e^{kt_{\text{now}}}.$$

Therefore our fossil's C_{14}/C_{12} ratio at any other time t is/was

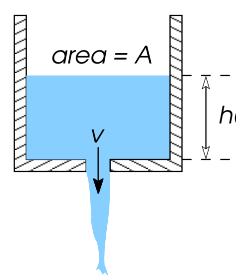
$$y(t) = y_{\text{now}}e^{k(t_{\text{now}}-t)}.$$

This allows us to compute the time at which the fossil died. At this time the C_{14}/C_{12} ratio must have been the common value in all living things, which can be measured – let's call it y_{life}. Then at the time t_{demise} when our fossil became a fossil we would have had $y(t_{\text{demise}}) = y_{\text{life}}$. Hence the age of the fossil would be given by

$$y_{\text{life}} = y(t_{\text{demise}}) = y_{\text{now}}e^{k(t_{\text{now}}-t_{\text{demise}})} \implies \boxed{t_{\text{now}} - t_{\text{demise}} = \frac{1}{k}\ln\frac{x_{\text{life}}}{x_{\text{now}}}}$$

10.2. Example: dating a leaky bucket. A bucket is filled with water. There is a hole in the bottom of the bucket so the water streams out at a certain rate.

$h(t)$	the height of water in the bucket
A	area of cross section of bucket
a	area of hole in the bucket
v	velocity with which water goes through the hole.



We have the following facts to work with:

- The amount (volume) of water in the bucket is $A \cdot h(t)$;
- The rate at which water is leaving the bucket is $a \cdot v(t)$;

Hence

$$\frac{\mathrm{d}Ah(t)}{\mathrm{d}t} = -av(t).$$

In fluid mechanics it is shown that the velocity of the water as it passes through the hole only depends on the height $h(t)$ of the water, and that, for some constant K,

$$v(t) = \sqrt{Kh(t)}.$$

The last two equations together give a differential equation for $h(t)$, namely,

$$\frac{\mathrm{d}h(t)}{\mathrm{d}t} = -\frac{a}{A}\sqrt{Kh(t)}.$$

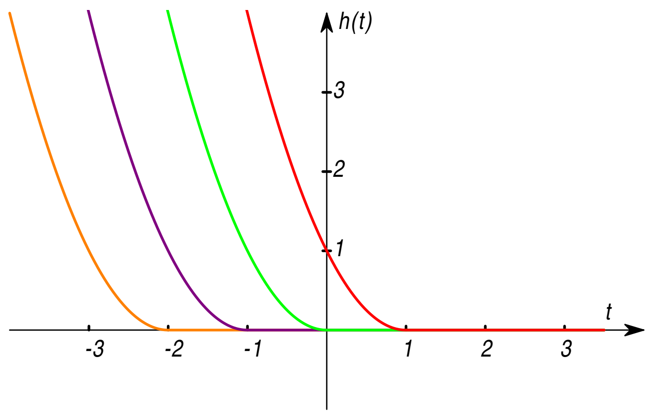


Figure 3. Several solutions $h(t; C)$ of the Leaking Bucket Equation (39). Note how they all have the same values when $t \geq 1$.

To make things a bit easier we assume that the constants are such that $\frac{a}{A}\sqrt{K} = 2$. Then $h(t)$ satisfies

$$h'(t) = -2\sqrt{h(t)}. \tag{39}$$

This equation is separable, and when we solve it we get

$$\frac{\mathrm{d}h}{2\sqrt{h}} = -1 \implies \sqrt{h(t)} = -t + C.$$

This formula can't be valid for *all* values of t, for if we take $t > C$, the RHS becomes negative and can't be equal to the square root in the LHS. But when $t \leq C$ we do get a solution,

$$h(t; C) = (C - t)^2.$$

This solution describes a bucket that is losing water until at time C it is empty. Motivated by the physical interpretation of our solution it is natural to assume that the bucket stays empty when $t > C$, so that the solution with integration constant C is given by

$$h(t) = \begin{cases} (C-t)^2 & \text{when } t \leq C \\ 0 & \text{for } t > C. \end{cases} \tag{40}$$

The snag appears again. (See § 3.2 and § 3.4.) Note that we had to divide by $\sqrt{h}$ to find the solution. This is not allowed when $h = 0$. It turns out that $h(t) = 0$ is a solution to the differential equation. The solution $h(t) = 0$ satisfies $h(0) = 0$, and our experience with differential equations so far would lead us to believe that this should therefore be the only solution that satisfies $h(0) = 0$. However, every solution from (40) with $C \leq 0$ also satisfies $h(0) = 0$. This problem is therefore different from the differential equations we have dealt with up to now. Namely, prescribing the value $y(0) = 0$ does not single out one solution of the differential equation (39).

10.3. Heat transfer. We all know that heat flows from warm to cold: if we put a cold spoon in a hot cup of coffee, then heat will flow from the coffee to the spoon. How fast does the spoon warm up?

According to physics the rate of change of the spoon's temperature is proportional to the difference in temperature between the coffee and the spoon. So, if T_c and T_s are the temperature of the coffee and the spoon, respectively, then

$$\frac{\mathrm{d}T_s}{\mathrm{d}t} = -K\big(T_s - T_c\big). \tag{41}$$

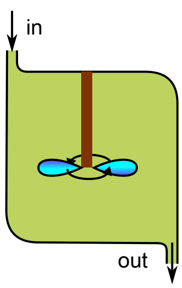
Here K is a constant that depends on the shape and material of the spoon, how much of the spoon is in the coffee, etc., but not on the temperatures T_s and T_c. If we assume that the spoon is small, then whatever small amount of heat it extracts from the coffee will not change the coffee temperature. Under that assumption we may assume that T_c is a constant and the differential equation (41) is both separable and linear so that we have two methods for solving it.

If the coffee itself is also cooling or warming up then T_c will depend on time and the equation (41) becomes

$$\frac{\mathrm{d}T_s}{\mathrm{d}t} + KT_s = KT_c(t). \tag{42}$$

If we know $T_c(t)$ then this is still a linear differential equation for T_s and we can solve it.

10.4. Mixing problems. Consider a container containing water and vinegar. If water and vinegar are flowing in and out of the container, then the concentration of vinegar in the container will change with time according to some differential equation. *Which* differential equation describes the vinegar content of the container depends on the precise details of the set-up.



As an example, let us assume that the container has fixed volume $V = 1000$ liters. This means that the amount of liquid that flows in during any time interval must equal the amount of liquid flowing out of the container (the liquid is "incompressible," i.e. its density is fixed.)

Suppose furthermore that a mixture of 5% vinegar and 95% water flows into the container at 40 liters per minute. And suppose also that the liquid in the container is thoroughly mixed, so that the liquid that flows out of the container has the same vinegar/water mixture as the entire container.

Problem: Let $D(t)$ be the fraction of the liquid in the container that is vinegar. How does $D(t)$ change with time?

Solution: Instead of tracking the concentration $D(t)$ of vinegar we will look at the total amount of vinegar in the container. This amount is $D(t)V$.

To find the differential equation we consider how much vinegar flows in and out of the container during a short time interval of length Δt (i.e. between time t and $t + \Delta t$):

- **in**: The liquid volume flowing into the tank during a time Δt is $40\Delta t$ liters. Since the vinegar concentration of the in-flowing liquid is 5%, this means that $5\% \cdot 40\Delta t = 2\Delta t$ liters of vinegar enter the container.
- **out**: Since $40\Delta t$ liters of liquid enter the container, the same amount of liquid must also leave the container. The liquid that leaves is the well-stirred mixture, and thus it contains $D(t) \cdot 40\Delta t$ liters vinegar.

In total we see that the change in vinegar content during a time Δt is

$$\Delta\big(\text{vinegar content}\big) = 2\Delta t - 40D(t)\Delta t. \tag{43}$$

To find the change in concentration we divide by the volume of the container

$$\Delta D = \frac{2\Delta t - 40D(t)\Delta t}{1000} = \frac{\Delta t}{500} - \frac{D(t)}{25}\Delta t.$$

We find the rate of change by dividing by Δt and letting $\Delta t \to 0$:

$$\frac{\mathrm{d}D}{\mathrm{d}t} = \frac{1}{500} - \frac{1}{25}D. \tag{44}$$

This equation is again both linear and separable so we have two methods for solving it.

11. Problems

1. Read Example 10.2 on "Leaky bucket dating" again. In that example we assumed that $\frac{a}{A}\sqrt{K} = 2$.

(a) Solve diffeq for $h(t)$ without assuming $\frac{a}{A}\sqrt{K} = 2$. Abbreviate $C = \frac{a}{A}\sqrt{K}$.

(b) If in an experiment one found that the bucket empties in 20 seconds after being filled to height 20 cm, then how much is the constant C?

2. A population of bacteria grows at a rate proportional to its size. Write and solve a differential equation that expresses this. If there are 1000 bacteria after one hour and 2000 bacteria after two hours, how many bacteria are there after three hours?

3. Rabbits in Madison have a birth rate of 5% per year and a death rate (from old age) of 2% per year. Each year 1000 rabbits get run over and 700 rabbits move in from Sun Prairie.

(a) Write a differential equation that describes Madison's rabbit population at time t.

(b) If there were 12,000 rabbits in Madison in 1991, how many are there in 1994?

4. [Group Problem] Consider a cup of soup that is cooling by exchanging heat with the room. If the temperature of the soup is $T_s(t)$, and if the temperature of the room is T_r, then Newton's cooling law claims that

$$\frac{\mathrm{d}T_s}{\mathrm{d}t} = -K(T_s - T_r)$$

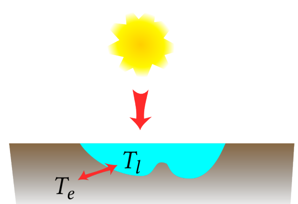
for some constant $K > 0$.

(a) What units does K have, if we measure temperature in degrees Fahrenheit, and time in seconds?

(b) Find the general solution to Newton's cooling law. What units does the arbitrary constant in your solution have? What is the limit of the temperature as $t \to \infty$?

(c) The soup starts at 180^oF, and sits in a room whose temperature is 75°F. In five minutes its temperature has dropped to 150°F. Find the cooling constant K. When will its temperature be 90°F?

5. [Group Problem] A lake gets heated by day when the sun is out, and loses heat at night. It also exchanges heat with the Earth. The Earth's temperature is T_e, and the lake's temperature is $T_l(t)$.



These effects together can be modeled by a differential equation for $T_l(t)$

$$\frac{\mathrm{d}T_l}{\mathrm{d}t} = -K(T_l - T_e) + S\sin(2\pi t).$$

Here the first term on the right represents Newton's cooling law, while the second term accounts for the heating and cooling by radiation: if $\sin(2\pi t) > 0$ then it's day and the sun is heating the lake, if $\sin(2\pi t) < 0$ then it's night and the lake is radiating heat into the cold sky. Time t is measured in days. Temperatures are in degrees Celsius.

(a) Assuming $K = S = 1$, and $T_e = 10$ find the general solution to the differential equation.

(b) Draw the direction field for the differential equation when $K = S = 1$ and $T_e = 10$. (First try to do this by hand. Then use a computer.)

(c) Does $\lim_{t\to\infty} T_l(t)$ exist? Consider the separate terms in the general solution you found in part **(a)**, and see if any of these terms have a limit as $t \to \infty$.

(d) Find the solution for arbitrary K and S (you may assume K and S are positive.)

6. Retaw is a mysterious living liquid; it grows at a rate of 5% of its volume per hour. A scientist has a tank initially holding y_0 gallons of retaw and removes retaw from the tank continuously at the rate of 3 gallons per hour.

(a) Find a differential equation for the number $y(t)$ of gallons of retaw in the tank at time t.

(b) Solve this equation for y as a function of t. (The initial volume y_0 will appear in our answer.)

(c) What is $\lim_{t\to\infty} y(t)$ if $y_0 = 100$?

(d) What should the value of y_0 be so that $y(t)$ remains constant?

7. A 1000 gallon vat is full of 25% solution of acid. Starting at time $t = 0$ a 40% solution of acid is pumped into the vat at 20 gallons per minute. The solution is kept well mixed and drawn off at 20 gallons per minute so as to maintain the total value of 1000 gallons. Derive an expression for the acid concentration at times $t > 0$. As $t \to \infty$ what percentage solution is approached?

8. [**Mixing**] The volume of a lake is $V = 10^9$ cubic feet. Pollution P runs into the lake at 3 cubic feet per minute, and clean water runs in at 21 cubic feet per minute. The lake drains at a rate of 24 cubic feet per minute so its volume is constant. Let C be the concentration of pollution in the lake; i.e. $C = P/V$.

(a) Give a differential equation for C.

(b) Solve the differential equation. Use the initial condition $C = C_0$ when $t = 0$ to evaluate the constant of integration.

(c) There is a critical value C^* with the property that for any solution $C = C(t)$ we have

$$\lim_{t\to\infty} C = C^*.$$

Find C^*. If $C_0 = C^*$, what is $C(t)$?

9. [**Mixing**] A 300 gallon tank is full of milk containing 2% butterfat. Milk containing 1% butterfat is pumped in a 10 gallons per minute starting at 10:00 AM and the well mixed milk is drained off at 15 gallons per minute. What is the percent butterfat in the milk in the tank 5 minutes later at 10:05 AM? Hint: How much milk is in the tank at time t? How much butterfat is in the milk at time $t = 0$?

10. **[Group Problem]** A philanthropist endows a chair. This means that she donates an amount of money B_0 to the university. The university invests the money (it earns interest) and pays the salary of a professor. Denote the interest rate on the investment by r (e.g. if $r = .06$, then the investment earns interest at a rate of 6% per year) the salary of the professor by a (e.g. $a = \$50,000$ per year), and the balance in the investment account at time t by B.

(a) Give a differential equation for B.

(b) Solve the differential equation. Use the initial condition $B = B_0$ when $t = 0$ to evaluate the constant of integration.

(c) There is a critical value B^* with the property that (1) if $B_0 < B^*$, then there is a $t > 0$ with $B(t) = 0$ (i.e. the account runs out of money) while (2) if $B_0 > B^*$, then $\lim_{t\to\infty} B = \infty$. Find B^*.

(d) This problem is like the pollution problem except for the signs of r and a. Explain.

11. **[Group Problem]** A citizen pays social security taxes of a dollars per year for T_1 years, then retires, then receives payments of b dollars per year for T_2 years, then dies. The account which receives and dispenses the money earns interest at a rate of r% per year and has no money at time $t = 0$ and no money at the time $t = T_1 + T_2$ of death. Find two differential equations for the balance $B(t)$ at time t; one valid for $0 \le t \le T_1$, the other valid for $T_1 \le t \le T_1 + T_2$. Express the ratio b/a in terms of T_1, T_2, and r. Reasonable values for T_1, T_2, and r are $T_1 = 40$, $T_2 = 20$, and $r = 5\% = .05$. This model ignores inflation. Notice that $0 < dB/dt$ for $0 < t < T_1$, that $dB/dt < 0$ for $T_1 < t < T_1 + T_2$, and that the account earns interest *even for* $T_1 < t < T_1 + T_2$.

CHAPTER 4

Taylor's Formula

All continuous functions that vanish at $x = a$
are approximately equal at $x = a$,
but some are more approximately equal than others.

1. Taylor Polynomials

Suppose we need to do some computation with a complicated function $y = f(x)$, and suppose that the only values of x we care about are close to some constant $x = a$. Since polynomials are simpler than most other functions, we could then look for a polynomial $y = P(x)$ that somehow "matches" our function $y = f(x)$ for values of x close to a. And we could then replace our function f with the polynomial P, hoping that the error we make is not too big. *Which* polynomial we choose depends on when we think a polynomial "matches" a function. In this chapter we will say that a polynomial P of degree n matches a function f at $x = a$ ***if P has the same value and the same derivatives of order $1, 2, \ldots, n$ at $x = a$ as the function f***. The polynomial that matches a given function at some point $x = a$ is the Taylor polynomial of f. It is given by the following formula.

1.1. Definition. *The Taylor polynomial of a function $y = f(x)$ of degree n at a point a is the polynomial*

$$T_n^a f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n. \tag{45}$$

(Recall that $n! = 1 \cdot 2 \cdot 3 \cdots n$, and by definition $0! = 1$.

1.2. Theorem. *The Taylor polynomial has the following property: it is the only polynomial $P(x)$ of degree n whose value and whose derivatives of orders $1, 2, \ldots,$ and n are the same as those of f, i.e. it's the only polynomial of degree n for which*

$$P(a) = f(a), \quad P'(a) = f'(a), \quad P''(a) = f''(a), \quad \ldots, \quad P^{(n)}(a) = f^{(n)}(a)$$

holds.

Proof. We do the case $a = 0$, for simplicity. Let n be given, consider a polynomial $P(x)$ of degree n, say,

$$P(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots + a_n x^n,$$

and let's see what its derivatives look like. They are:

$$\begin{array}{rclllllllll}
P(x) & = & a_0 & + & a_1 x & + & a_2 x^2 & + & a_3 x^3 & + & a_4 x^4 & + \cdots \\
P'(x) & = & & & a_1 & + & 2a_2 x & + & 3a_3 x^2 & + & 4a_4 x^3 & + \cdots \\
P^{(2)}(x) & = & & & & & 1\cdot 2a_2 & + & 2\cdot 3a_3 x & + & 3\cdot 4a_4 x^2 & + \cdots \\
P^{(3)}(x) & = & & & & & & & 1\cdot 2\cdot 3a_3 & + & 2\cdot 3\cdot 4a_4 x & + \cdots \\
P^{(4)}(x) & = & & & & & & & & & 1\cdot 2\cdot 3\cdot 4a_4 & + \cdots
\end{array}$$

When we set $x = 0$ all the terms that have a positive power of x vanish, and we are left with the first entry on each line, i.e.

$$P(0) = a_0, \quad P'(0) = a_1, \quad P^{(2)}(0) = 2a_2, \quad P^{(3)}(0) = 2\cdot 3a_3, \text{ etc.}$$

and in general

$$P^{(k)}(0) = k!a_k \text{ for } 0 \le k \le n.$$

For $k \ge n+1$ the derivatives $P^{(k)}(x)$ all vanish of course, since $P(x)$ is a polynomial of degree n.

Therefore, if we want P to have the same values and derivatives at $x = 0$ of orders 1,,..., n as the function f, then we must have $k!a_k = P^{(k)}(0) = f^{(k)}(0)$ for all $k \le n$. Thus

$$a_k = \frac{f^{(k)}(0)}{k!} \qquad \text{for } 0 \le k \le n.$$

□

2. Examples

2.1. Taylor polynomials of order zero and one. The zeroth order Taylor polynomial of a function $f(x)$ is written as $T_0^a f(x)$. According to the definition (45) it is given by

$$T_0^a f(x) = f(a).$$

It does not depend on x and is just a constant.

The first order Taylor polynomial is

$$T_1^a f(x) = f(a) + f'(a)(x-a).$$

The graph of the function

$$y = T_1^a f(x), \text{ i.e. } y = f(a) + f'(a)(x-a),$$

is the tangent line at $x = a$ to the graph of the function $y = f(x)$. The function $y = T_1^a f(x)$ is exactly the *linear approximation of* $f(x)$ for x close to a that was derived in 1st semester calculus.

The Taylor polynomial generalizes this first order approximation by providing "higher order approximations" to f.

Most of the time we will take $a = 0$ in which case we write $T_n f(x)$ instead of $T_n^a f(x)$, and we get a slightly simpler formula

$$T_n f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \cdots + \frac{f^{(n)}(0)}{n!}x^n. \tag{46}$$

You will see below that for many functions $f(x)$ the Taylor polynomials $T_n f(x)$ give better and better approximations as we add more terms (i.e. as we increase n). For this reason the limit when $n \to \infty$ is often considered, which leads to the *infinite sum*

$$T_\infty f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \cdots$$

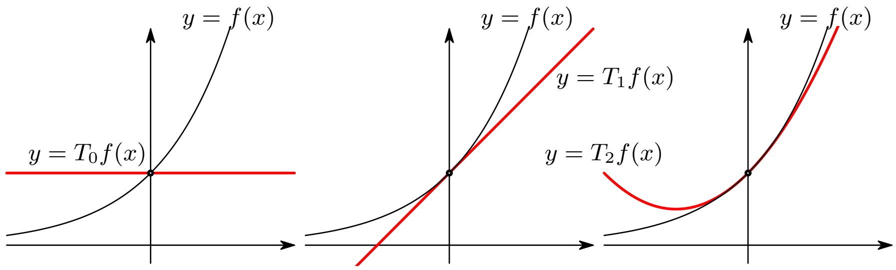


Figure 1. The Taylor polynomials of degree 0, 1 and 2 of $f(x) = e^x$ at $a = 0$. The zeroth order Taylor polynomial has the right value at $x = 0$ but it doesn't know whether or not the function f is increasing at $x = 0$. The first order Taylor polynomial has the right slope at $x = 0$, but it doesn't see if the graph of f is curved up or down at $x = 0$. The second order Taylor polynomial also has the right curvature at $x = 0$.

Although we will not try to make sense of the "sum of infinitely many numbers" at this point, we will return to this question in the next chapter on Sequences and Series.

2.2. Example: Compute the Taylor polynomials of degree 0, 1 and 2 of $f(x) = e^x$ at $a = 0$, and plot them. One has

$$f(x) = e^x \implies f'(x) = e^x \implies f''(x) = e^x,$$

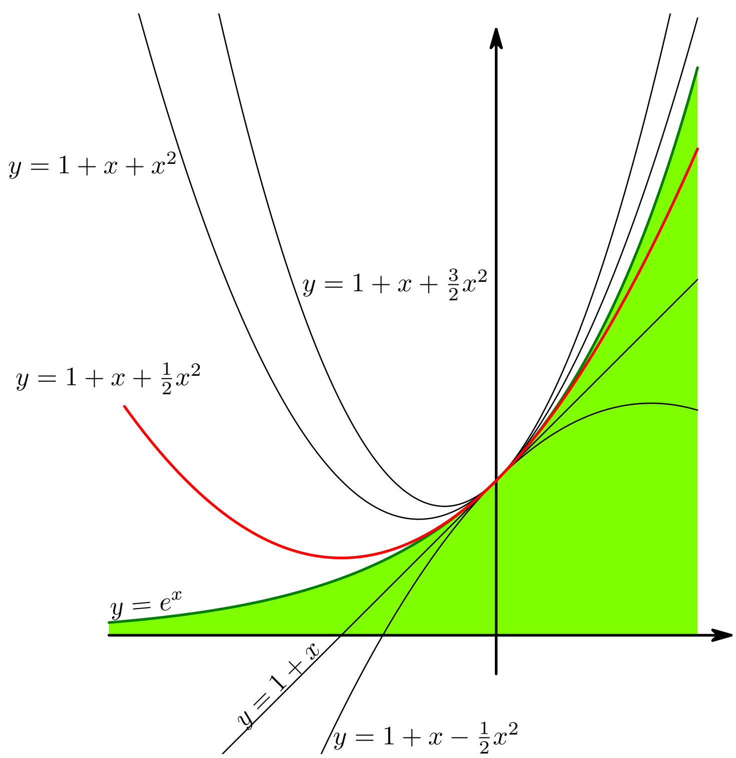


Figure 2. The top edge of the shaded region is the graph of $y = e^x$. The graphs are of the functions $y = 1 + x + Cx^2$ for various values of C. These graphs all are tangent at $x = 0$, but one of the parabolas matches the graph of $y = e^x$ better than any of the others.

so that

$$f(0) = 1, \quad f'(0) = 1, \quad f''(0) = 1.$$

Therefore the first three Taylor polynomials of e^x at $a = 0$ are

$$\begin{aligned} T_0 f(x) &= 1 \\ T_1 f(x) &= 1 + x \\ T_2 f(x) &= 1 + x + \frac{1}{2}x^2. \end{aligned}$$

The graphs are found in Figure 1.

The Taylor polynomial of degree 0, i.e. $T_0 f(x) = 1$ captures the fact that e^x by virtue of its continuity does not change very much if x stays close to $x = 0$.

The Taylor polynomial of degree 1, i.e. $T_1 f(x) = 1 + x$ corresponds to the tangent line to the graph of $f(x) = e^x$, and so it also captures the fact that the function $f(x)$ is increasing near $x = 0$.

Clearly $T_1 f(x)$ is a better approximation to e^x than $T_0 f(x)$.

The graphs of both $y = T_0 f(x)$ and $y = T_1 f(x)$ are straight lines, while the graph of $y = e^x$ is curved (in fact, convex). The second order Taylor polynomial captures this convexity. In fact, the graph of $y = T_2 f(x)$ is a parabola, and since it has the same first and second derivative at $x = 0$, its curvature is the same as the curvature of the graph of $y = e^x$ at $x = 0$. So it seems that $y = T_2 f(x) = 1 + x + x^2/2$ is an approximation to $y = e^x$ that beats both $T_0 f(x)$ and $T_1 f(x)$.

Figure 2 shows the graphs of various parabolas that have the same tangent line as the graph of $y = e^x$ at $x = 0$. Such parabolas are given by $y = 1 + x + Cx^2$, for arbitrary C. The figure shows that the choice $C = \frac{1}{2}$ leads to the parabola that best matches the graph of $y = e^x$.

2.3. Example: Find the Taylor polynomials of $f(x) = \sin x$. When we start computing the derivatives of $\sin x$ we find

$$f(x) = \sin x, \quad f'(x) = \cos x, \quad f''(x) = -\sin x, \quad f^{(3)}(x) = -\cos x,$$

and thus

$$f^{(4)}(x) = \sin x.$$

So after four derivatives we're back to where we started, and the sequence of derivatives of $\sin x$ cycles through the pattern

$$\sin x,\ \cos x,\ -\sin x,\ -\cos x,\ \sin x,\ \cos x,\ -\sin x,\ -\cos x,\ \sin x,\ \ldots$$

on and on. At $x = 0$ we then get the following values for the derivatives $f^{(j)}(0)$,

j	0	1	2	3	4	5	6	7	8	$\cdots$
$f^{(j)}(0)$	0	1	0	-1	0	1	0	-1	0	$\cdots$

This gives the following Taylor polynomials

$$
\begin{aligned}
T_0 f(x) &= 0 \\
T_1 f(x) &= x \\
T_2 f(x) &= x \\
T_3 f(x) &= x - \frac{x^3}{3!} \\
T_4 f(x) &= x - \frac{x^3}{3!} \\
T_5 f(x) &= x - \frac{x^3}{3!} + \frac{x^5}{5!}
\end{aligned}
$$

Note that since $f^{(2)}(0) = 0$ the Taylor polynomials $T_1 f(x)$ and $T_2 f(x)$ are the same! The second order Taylor polynomial in this example is really only a polynomial of degree one. In general the Taylor polynomial $T_n f(x)$ of any function is a polynomial of degree at most n, and this example shows that the degree can sometimes be strictly less.

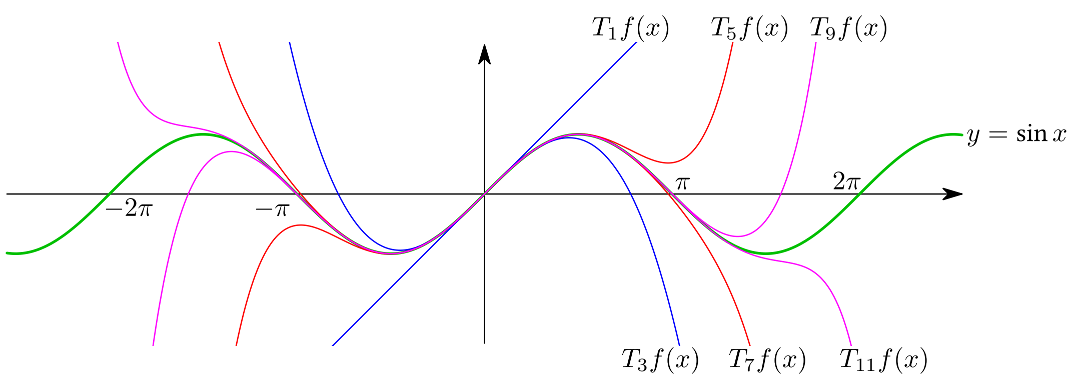


Figure 3. Taylor polynomials of $f(x) = \sin x$

2.4. Example: compute the Taylor polynomials of degree two and three of $f(x) = 1 + x + x^2 + x^3$ at $a = 3$.

Solution: Remember that our notation for the n^{th} degree Taylor polynomial of a function f at a is $T_n^a f(x)$, and that it is defined by (45).

We have

$$f'(x) = 1 + 2x + 3x^2, \quad f''(x) = 2 + 6x, \quad f'''(x) = 6$$

Therefore $f(3) = 40$, $f'(3) = 34$, $f''(3) = 20$, $f'''(3) = 6$, and thus

$$T_2^3 f(x) = 40 + 34(x-3) + \frac{20}{2!}(x-3)^2 = 40 + 34(x-3) + 10(x-3)^2. \tag{47}$$

Why don't we expand the answer? You could do this (i.e. replace $(x-3)^2$ by $x^2 - 6x + 9$ throughout and sort the powers of x), but as we will see in this chapter, the Taylor polynomial $T_n^a f(x)$ is used as an approximation for $f(x)$ when x is close to a. In this example $T_2^3 f(x)$ is to be used when x is close to 3. If $x - 3$ is a small number then the successive powers $x - 3$, $(x-3)^2$, $(x-3)^3$, …decrease rapidly, and so the terms in (47) are arranged in decreasing order.

We can also compute the third degree Taylor polynomial. It is

$$\begin{aligned} T_3^3 f(x) &= 40 + 34(x-3) + \frac{20}{2!}(x-3)^2 + \frac{6}{3!}(x-3)^3 \\ &= 40 + 34(x-3) + 10(x-3)^2 + (x-3)^3. \end{aligned}$$

If we expand this (this takes a little work) we find that

$$40 + 34(x-3) + 10(x-3)^2 + (x-3)^3 = 1 + x + x^2 + x^3.$$

So the third degree Taylor polynomial is the function f itself! Why is this so? Because of Theorem 1.2! Both sides in the above equation are third degree polynomials, and their derivatives of order 0, 1, 2 and 3 are the same at $x = 3$, so, since there is only one polynomial with this property, they must be the same polynomial.

3. Some special Taylor polynomials

Here is a list of functions whose Taylor polynomials are sufficiently regular that we can write a formula for the n^{th} term.

$$\begin{aligned} T_n e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} \\ T_{2n+1}\{\sin x\} &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} \\ T_{2n}\{\cos x\} &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots + (-1)^n \frac{x^{2n}}{(2n)!} \\ T_n\left\{\frac{1}{1-x}\right\} &= 1 + x + x^2 + x^3 + x^4 + \cdots + x^n \qquad \text{(Geometric Sum)} \\ T_n\{\ln(1+x)\} &= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots + (-1)^{n+1}\frac{x^n}{n} \end{aligned}$$

All of these Taylor polynomials can be computed directly from the definition, by repeatedly differentiating $f(x)$.

Another function whose Taylor polynomial we should know is $f(x) = (1+x)^a$, where a is a constant. You can compute $T_n f(x)$ directly from the definition, and when we do this we find

$$\begin{aligned} T_n\{(1+x)^a\} = 1 + ax + \frac{a(a-1)}{1\cdot 2}x^2 + \frac{a(a-1)(a-2)}{1\cdot 2\cdot 3}x^3 \\ + \cdots + \frac{a(a-1)\cdots(a-n+1)}{1\cdot 2\cdots n}x^n. \end{aligned} \tag{48}$$

Note that here a is not assumed to be an integer. This formula is called ***Newton's binomial formula***. The coefficient of x^n is called a ***binomial coefficient***, and it is written

$$\binom{a}{n} = \frac{a(a-1)\cdots(a-n+1)}{n!}. \tag{49}$$

When a is an integer $\binom{a}{n}$ is also called "a choose n." Using this notation equation (48) can be written as

$$T_n\{(1+x)^a\} = 1 + \binom{a}{1}x + \binom{a}{2}x^2 + \binom{a}{3}x^3 + \cdots + \binom{a}{n}x^n.$$

Note that we already knew special cases of the binomial formula: when a is a positive integer the binomial coefficients are just the numbers in ***Pascal's triangle***.

4. Problems

1. Find a second order polynomial (i.e. a quadratic function) $Q(x)$ such that $Q(7) = 43, Q'(7) = 19, Q''(7) = 11$.

2. Find a second order polynomial $p(x)$ such that $p(2) = 3$, $p'(2) = 8$, and $p''(2) = -1$.

3. A Third order polynomial $P(x)$ satisfies $P(0) = 1, P'(0) = -3, P''(0) = -8, P'''(0) = 24$. Find $P(x)$.

4. Let $f(x) = \sqrt{x+25}$. Find the polynomial $P(x)$ of degree three such that $P^{(k)}(0) = f^{(k)}(0)$ for $k = 0, 1, 2, 3$.

5. Let $f(x) = 1 + x - x^2 - x^3$. Compute and graph $T_0 f(x)$, $T_1 f(x)$, $T_2 f(x)$, $T_3 f(x)$, and $T_4 f(x)$, as well as $f(x)$ itself (so, for each of these functions find where they are positive or negative, where they are increasing/decreasing, and find the inflection points on their graph.)

6. **(a)** Find $T_3 \sin x$ and $T_5 \sin x$.

(b) Graph $T_3 \sin x$ and $T_5 \sin x$ as well as $y = \sin x$ in one picture. (As before, find where these functions are positive or negative, where they are increasing/decreasing, and find the inflection points on their graph. This problem can&should be done without a graphing calculator.)

* * *

Compute $T_0^a f(x)$, $T_1^a f(x)$ and $T_2^a f(x)$ for the following functions.

7. $f(x) = x^3$, $a = 0$; then for $a = 1$ and $a = 2$.

8. $f(x) = \dfrac{1}{x}$, $a = 1$. Also do $a = 2$.

9. $f(x) = \sqrt{x}$, $a = 1$.

10. $f(x) = \ln x$, $a = 1$. Also $a = e^2$.

11. $f(x) = \ln \sqrt{x}$, $a = 1$.

12. $f(x) = \sin(2x)$, $a = 0$, also $a = \pi/4$.

13. $f(x) = \cos(x)$, $a = \pi$.

14. $f(x) = (x-1)^2$, $a = 0$, and also $a = 1$.

15. $f(x) = \dfrac{1}{e^x}$, $a = 0$.

16. Find the nth degree Taylor polynomial $T_n^a f(x)$ of the following functions $f(x)$

n	a	$f(x)$
2	0	$1 + x - x^3$
3	0	$1 + x - x^3$
25	0	$1 + x - x^3$
25	2	$1 + x - x^3$
2	1	$1 + x - x^3$
1	1	x^2
2	1	x^2
5	1	$1/x$
5	0	$1/(1+x)$
5	1	$1/(1+x)$
5	$-\frac{1}{2}$	$1/(1+x)$
3	0	$1/(1-3x+2x^2)$

For which of these combinations $(n, a, f(x))$ is $T_n^a f(x)$ the same as $f(x)$?

* * *

Compute the Taylor series $T_\infty f(t)$ for the following functions (α is a constant). Give a formula for the coefficient of x^n in $T_\infty f(t)$. (*Be smart. Remember properties of the logarithm, definitions of the hyperbolic functions, partial fraction decomposition.*)

17. e^t

18. $e^{\alpha t}$

19. $\sin(3t)$

20. $\sinh t$

21. $\cosh t$

22. $\dfrac{1}{1+2t}$

23. $\dfrac{3}{(2-t)^2}$

24. $\ln(1+t)$

25. $\ln(2+2t)$

26. $\ln \sqrt{1+t}$

27. $\ln(1+2t)$

28. $\ln \sqrt{\dfrac{1+t}{1-t}}$

29. $\dfrac{1}{1-t^2}$ [hint:PFD!]

30. $\dfrac{t}{1-t^2}$

31. $\sin t + \cos t$

32. $2 \sin t \cos t$

33. $\tan t$ (3 terms only)

34. $1 + t^2 - \dfrac{2}{3}t^4$

35. $(1 + t)^5$

36. $\sqrt[3]{1 + t}$

37. $f(x) = \frac{x^4}{1+4x^2}$, what is $f^{(10)}(0)$?

38. **[Group Problem]** Compute the Taylor series of the following two functions

$$f(x) = \sin a \cos x + \cos a \sin x$$

and

$$g(x) = \sin(a + x)$$

where a is a constant.

39. **[Group Problem]** Compute the Taylor series of the following two functions

$$h(x) = \cos a \cos x - \sin a \sin x$$

and

$$k(x) = \cos(a + x)$$

where a is a constant.

40. **[Group Problem]** The following questions ask us to rediscover *Newton's Binomial Formula*, which is just the Taylor series for $(1 + x)^n$. Newton's formula generalizes the formulas for $(a + b)^2$, $(a + b)^3$, etc that we get using Pascal's triangle. It allows non integer exponents which are allowed to be either positive and negative. Reread section 3 before doing this problem.

(a) Find the Taylor series of $f(x) = \sqrt{1 + x}$ $(= (1 + x)^{1/2})$

(b) Find the coefficient of x^4 in the Taylor series of $f(x) = (1+x)^\pi$ (don't do the arithmetic!)

(c) Let p be any real number. Compute the terms of degree 0, 1, 2 and 3 of the Taylor series of

$$f(x) = (1 + x)^p$$

(d) Compute the Taylor polynomial of degree n of $f(x) = (1 + x)^p$.

(e) Write the result of (d) for the exponents $p = 2, 3$ and also, for $p = -1, -2, -3$ and finally for $p = \frac{1}{2}$. The *Binomial Theorem* states that this series converges when $|x| < 1$.

5. The Remainder Term

The Taylor polynomial $T_n f(x)$ is almost never exactly equal to $f(x)$, but often it is a good approximation, especially if x is small. To see how good the approximation is we define the "error term" or, "remainder term."

5.1. Definition. *If f is an n times differentiable function on some interval containing a, then*

$$R_n^a f(x) = f(x) - T_n^a f(x)$$

is called the n^{th} order remainder (or error) term in the Taylor expansion of f.

If $a = 0$, as will be the case in most examples we do, then we write

$$R_n f(x) = f(x) - T_n f(x).$$

These definitions let us write any function $f(x)$ as "Taylor polynomial plus error," i.e.

$$f(x) = \underbrace{T_n^a f(x)}_{\text{approximation}} + \underbrace{R_n^a f(x)}_{\text{the error}} \tag{50}$$

Generally the approximation is something we know how to compute, while we can only hope to prove that the remainder is in some sense "small."

5.2. Example. If $f(x) = \sin x$ then we have found that $T_3 f(x) = x - \frac{1}{6}x^3$, so that

$$R_3\{\sin x\} = \sin x - x + \tfrac{1}{6}x^3.$$

This is a completely correct formula for the remainder term, but it is rather useless: there is nothing about this expression that suggests that $x-\frac{1}{6}x^3$ is a much better approximation to $\sin x$ than, say, $x+\frac{1}{6}x^3$.

The usual situation is that there is no simple formula for the remainder term.

5.3. An unusual example, in which there *is* a simple formula for $R_n f(x)$. Consider $f(x) = 1 - x + 3\,x^2 - 15\,x^3$. Then we find

$$T_2 f(x) = 1 - x + 3\,x^2, \text{ so that } R_2 f(x) = f(x) - T_2 f(x) = -15\,x^3.$$

Thus

$$f(x) = \underbrace{1 - x + 3\,x^2}_{\text{approximation}} - \underbrace{15\,x^3}_{\text{error}}$$

The moral of this example is this: *Given a polynomial $f(x)$ we find its n^{th} degree Taylor polynomial at $a = 0$ by taking all terms of degree $\le n$ in $f(x)$; the remainder $R_n f(x)$ then consists of the remaining terms.*

5.4. Another unusual, but important example where we can compute $R_n f(x)$. Consider the function

$$f(x) = \frac{1}{1-x}.$$

Then repeated differentiation gives

$$f'(x) = \frac{1}{(1-x)^2}, \quad f^{(2)}(x) = \frac{1\cdot 2}{(1-x)^3}, \quad f^{(3)}(x) = \frac{1\cdot 2\cdot 3}{(1-x)^4}, \quad \ldots$$

and thus

$$f^{(n)}(x) = \frac{1\cdot 2\cdot 3\cdots n}{(1-x)^{n+1}}.$$

Consequently,

$$f^{(n)}(0) = n! \implies \frac{1}{n!}f^{(n)}(0) = 1,$$

and we see that the Taylor polynomials of this function are really simple, namely

$$T_n f(x) = 1 + x + x^2 + x^3 + x^4 + \cdots + x^n.$$

But this sum should be really familiar: it is just the ***Geometric Sum*** (each term is x times the previous term). Its sum is given by[1]

$$T_n f(x) = 1 + x + x^2 + x^3 + x^4 + \cdots + x^n = \frac{1 - x^{n+1}}{1-x},$$

which we can rewrite as

$$T_n f(x) = \frac{1}{1-x} - \frac{x^{n+1}}{1-x} = f(x) - \frac{x^{n+1}}{1-x}.$$

The remainder term therefore is

$$R_n f(x) = f(x) - T_n f(x) = \frac{x^{n+1}}{1-x}.$$

[1]Multiply both sides with $1 - x$ to verify this, in case we had forgotten the formula!

6. Lagrange's Formula for the Remainder Term

6.1. Theorem. *Let f be an $n+1$ times differentiable function on some interval I containing $x=0$. Then for every x in the interval I there is a ξ between 0 and x such that*

$$R_n f(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}x^{n+1}.$$

(ξ between 0 and x means either $0<\xi<x$ or $x<\xi<0$, depending on the sign of x.)

This theorem (including the proof) is similar to the Mean Value Theorem. The proof is a bit involved, and is given at the end of this chapter.

There are calculus textbooks that, after presenting this remainder formula, give a whole bunch of problems that ask us to find ξ for given f and x. Such problems completely miss the point of Lagrange's formula. The point is that *even though we usually can't compute the mystery point ξ precisely, Lagrange's formula for the remainder term allows us to **estimate** the remainder.* Here is the most common way to estimate the remainder:

6.2. Estimate of remainder term. *If f is an $n+1$ times differentiable function on an interval containing $x=0$, and if we have a constant M such that*

(†) $$\left|f^{(n+1)}(t)\right| \le M \text{ for all } t \text{ between } 0 \text{ and } x,$$

then

$$|R_n f(x)| \le \frac{M|x|^{n+1}}{(n+1)!}.$$

Proof. We don't know what ξ is in Lagrange's formula, but it doesn't matter, for wherever it is, it must lie between 0 and x so that our assumption (†) implies $|f^{(n+1)}(\xi)| \le M$. Put that in Lagrange's formula and we get the stated inequality. □

6.3. How to compute e in a few decimal places. We are used to being able to find decimal approximations to numbers such as e (or $e^{0.315\,219}$) by relying on the magic of an electronic calculator or computer. How does a calculator, which in principle only knows how to add, subtract, multiply, and divide numbers, compute such numbers? One way is to use the Taylor expansion with Lagrange's remainder term. In this section we will see how to compute e itself in six decimals.

Consider $f(x)=e^x$. We computed the Taylor polynomials before. If we set $x=1$, then we get $e=f(1)=T_n f(1)+R_n f(1)$, and thus, taking $n=8$,

$$e = 1+\frac{1}{1!}+\frac{1}{2!}+\frac{1}{3!}+\frac{1}{4!}+\frac{1}{5!}+\frac{1}{6!}+\frac{1}{7!}+\frac{1}{8!}+R_8(1).$$

By Lagrange's formula there is a ξ between 0 and 1 such that

$$R_8(1) = \frac{f^{(9)}(\xi)}{9!}\,1^9 = \frac{e^\xi}{9!}.$$

(remember: $f(x)=e^x$, so all its derivatives are also e^x.) We don't really know where ξ is, but since it lies between 0 and 1 we know that $1<e^\xi<e$. So the remainder term $R_8(1)$ is positive and no more than $e/9!$. Estimating $e<3$, we find

$$\frac{1}{9!} < R_8(1) < \frac{3}{9!}.$$

Thus we see that

$$1+\frac{1}{1!}+\frac{1}{2!}+\frac{1}{3!}+\cdots+\frac{1}{7!}+\frac{1}{8!}+\frac{1}{9!} < e < 1+\frac{1}{1!}+\frac{1}{2!}+\frac{1}{3!}+\cdots+\frac{1}{7!}+\frac{1}{8!}+\frac{3}{9!}$$

or, in decimals,

$$2.718\,281\ldots < e < 2.718\,287\ldots$$

6.4. Error in the approximation $\sin x \approx x$. In many calculations involving $\sin x$ for small values of x one makes the simplifying approximation $\sin x \approx x$, justified by the known limit

$$\lim_{x\to 0} \frac{\sin x}{x} = 1.$$

6.4.1. *Question: how big is the error in this approximation?* To answer this question, we use Lagrange's formula for the remainder term again.

Let $f(x) = \sin x$. Then the first degree Taylor polynomial of f is

$$T_1 f(x) = x.$$

The approximation $\sin x \approx x$ is therefore exactly what we get if we approximate $f(x) = \sin x$ by its first degree Taylor polynomial. Lagrange tells us that

$$f(x) = T_1 f(x) + R_1 f(x), \text{ i.e. } \sin x = x + R_1 f(x),$$

where, since $f''(x) = -\sin x$,

$$R_1 f(x) = \frac{f''(\xi)}{2!} x^2 = -\tfrac{1}{2} \sin\xi \cdot x^2$$

for some ξ between 0 and x.

As always with Lagrange's remainder term, we don't know where ξ is precisely, so we have to estimate the remainder term. The easiest way to do this (but not the best: see below) is to say that no matter what ξ is, $\sin\xi$ will always be between -1 and 1. Hence the remainder term is bounded by

$$|R_1 f(x)| \le \tfrac{1}{2} x^2, \tag{¶}$$

and we find that

$$x - \tfrac{1}{2}x^2 \le \sin x \le x + \tfrac{1}{2}x^2.$$

6.4.2. *Question: How small must we choose x to be sure that the approximation $\sin x \approx x$ isn't off by more than 1% ?* If we want the error to be less than 1% of the estimate, then we should require $\frac{1}{2}x^2$ to be less than 1% of $|x|$, i.e.

$$\tfrac{1}{2}x^2 < 0.01 \cdot |x| \Leftrightarrow |x| < 0.02$$

So we have shown that, if we choose $|x| < 0.02$, then the error we make in approximating $\sin x$ by just x is no more than 1%.

A final comment about this example: the estimate for the error we got here can be improved quite a bit in two different ways:

(1) You could notice that one has $|\sin x| \le x$ for all x, so if ξ is between 0 and x, then $|\sin\xi| \le |\xi| \le |x|$, which gives us the estimate

$$|R_1 f(x)| \le \tfrac{1}{2}|x|^3 \quad \text{instead of } \tfrac{1}{2}x^2 \text{ as in (¶)}.$$

(2) For this particular function the two Taylor polynomials $T_1 f(x)$ and $T_2 f(x)$ are the same (because $f''(0) = 0$). So $T_2 f(x) = x$, and we can write

$$\sin x = f(x) = x + R_2 f(x),$$

In other words, the error in the approximation $\sin x \approx x$ is also given by the *second* order remainder term, which according to Lagrange is given by

$$R_2 f(x) = \frac{-\cos\xi}{3!} x^3 \quad \overset{|\cos\xi|\le 1}{\Longrightarrow} \quad |R_2 f(x)| \le \tfrac{1}{6}|x|^3,$$

which is the best estimate for the error in $\sin x \approx x$ we have so far.

7. Problems

1. Find the fourth degree Taylor polynomial $T_4\{\cos x\}$ for the function $f(x) = \cos x$ and estimate the error $|\cos x - T_4\{\cos x\}|$ for $|x| < 1$.

2. Find the 4th degree Taylor polynomial $T_4\{\sin x\}$ for the function $f(x) = \sin x$. Estimate the error $|\sin x - T_4\{\sin x\}|$ for $|x| < 1$.

3. (*Computing the cube root of* 9) The cube root of $8 = 2 \times 2 \times 2$ is easy, and 9 is only one more than 8. So we could try to compute $\sqrt[3]{9}$ by viewing it as $\sqrt[3]{8+1}$.

(a) Let $f(x) = \sqrt[3]{8+x}$. Find $T_2 f(x)$, and estimate the error $|\sqrt[3]{9} - T_2 f(1)|$.

(b) Repeat part (**a**) for "$n = 3$", i.e. compute $T_3 f(x)$ and estimate $|\sqrt[3]{9} - T_3 f(1)|$.

4. Follow the method of problem 7.3 to compute $\sqrt{10}$:

(a) Use Taylor's formula with $f(x) = \sqrt{9+x}$, $n = 1$, to calculate $\sqrt{10}$ approximately. Show that the error is less than $1/216$.

(b) Repeat with $n = 2$. Show that the error is less than 0.0003.

5. Find the eighth degree Taylor polynomial $T_8 f(x)$ about the point 0 for the function $f(x) = \cos x$ and estimate the error $|\cos x - T_8 f(x)|$ for $|x| < 1$.

Next, find the ninth degree Taylor polynomial, and estimate $|\cos x - T_9 f(x)|$ for $|x| \leq 1$.

8. The limit as $x \to 0$, keeping n fixed

8.1. Little-oh. Lagrange's formula for the remainder term lets us write a function $y = f(x)$ that is defined on some interval containing $x = 0$, in the following way

$$f(x) = f(0) + f'(0)x + \frac{f^{(2)}(0)}{2!}x^2 + \cdots + \frac{f^{(n)}(0)}{n!}x^n + \frac{f^{(n+1)}(\xi)}{(n+1)!}x^{n+1} \tag{51}$$

The last term contains the ξ from Lagrange's theorem, which depends on x, and of which we only know that it lies between 0 and x. For many purposes it is not necessary to know the last term in this much detail – often it is enough to know that "in some sense" the last term is the smallest term, in particular, as $x \to 0$ it is much smaller than x, or x^2, or, ..., or x^n:

8.2. Theorem. *If the $n+1$st derivative $f^{(n+1)}(x)$ is continuous at $x = 0$ then the remainder term $R_n f(x) = f^{(n+1)}(\xi)x^{n+1}/(n+1)!$ satisfies*

$$\lim_{x\to 0} \frac{R_n f(x)}{x^k} = 0$$

for any $k = 0, 1, 2, \ldots, n$.

Proof. Since ξ lies between 0 and x, one has $\lim_{x\to 0} f^{(n+1)}(\xi) = f^{(n+1)}(0)$, and therefore

$$\lim_{x\to 0} \frac{R_n f(x)}{x^k} = \lim_{x\to 0} f^{(n+1)}(\xi)\frac{x^{n+1}}{x^k} = \lim_{x\to 0} f^{(n+1)}(\xi) \cdot x^{n+1-k} = f^{(n+1)}(0) \cdot 0 = 0.$$

$\square$

So we can rephrase (51) by saying

$$f(x) = f(0) + f'(0)x + \frac{f^{(2)}(0)}{2!}x^2 + \cdots + \frac{f^{(n)}(0)}{n!}x^n + \text{ remainder}$$

where the remainder is much smaller than x^n, x^{n-1}, ..., x^2, x or 1. In order to express the condition that some function is "much smaller than x^n," at least for very small x, Landau introduced the following notation which many people find useful.

8.3. Definition. *"$o(x^n)$" is an abbreviation for* ***any*** *function $h(x)$ that satisfies*

$$\lim_{x\to 0} \frac{h(x)}{x^n} = 0.$$

One interpretation of "$h(x) = o(x^n)$" is that the function $h(x)$ vanishes as $x \to 0$, and that it goes to zero "faster than x^n."

The definition says that $o(x^n)$ refers to ***any*** function with the indicated property. This means that different instances of $o(x^n)$ in a formula may refer to different functions of x (just as different $+C$'s in an integration may refer to different constants.) This makes computations with little-oh a bit different from the normal algebra that we are used to, as we will explain below (§ 8.6). Nevertheless, once we have understood this particular point, computations with little-oh are much simpler than with the Lagrange remainder term.

With the little-oh notation we can rewrite (51) as

$$f(x) = f(0) + f'(0)x + \frac{f^{(2)}(0)}{2!}x^2 + \cdots + \frac{f^{(n)}(0)}{n!}x^n + o(x^n).$$

The nice thing about Landau's little-oh is that we can compute with it, as long as we obey the following (at first sight rather strange) rules that will be proved in class

$$\begin{aligned} x^n \cdot o(x^m) &= o(x^{n+m}) && \\ o(x^n) \cdot o(x^m) &= o(x^{n+m}) && \\ x^m &= o(x^n) && \text{if } n < m \\ o(x^n) + o(x^m) &= o(x^n) && \text{if } n < m \\ o(Cx^n) &= o(x^n) && \text{for any constant } C \end{aligned}$$

8.4. Example: prove one of these little-oh rules. Let's do the first one, i.e. let's show that $x^n \cdot o(x^m)$ is $o(x^{n+m})$ as $x \to 0$.

Remember, if someone writes $x^n \cdot o(x^m)$, then the $o(x^m)$ is an abbreviation for some function $h(x)$ that satisfies $\lim_{x\to 0} h(x)/x^m = 0$. So the $x^n \cdot o(x^m)$ we are given here really is an abbreviation for $x^n h(x)$. We then have

$$\lim_{x\to 0} \frac{x^n h(x)}{x^{n+m}} = \lim_{x\to 0} \frac{h(x)}{x^m} = 0, \text{ since } h(x) = o(x^m).$$

8.5. Can we see that $x^3 = o(x^2)$ by looking at the graphs of these functions? A picture is of course never a proof, but have a look at figure 4 which shows us the graphs of $y = x$, x^2, x^3, x^4, x^5 and x^{10}. As we see, when x approaches 0, the graphs of higher powers of x approach the x-axis (much?) faster than do the graphs of lower powers.

You should also have a look at figure 5 which exhibits the graphs of $y = x^2$, as well as several linear functions $y = Cx$ (with $C = 1, \frac{1}{2}, \frac{1}{5}$ and $\frac{1}{10}$.) For each of these linear functions one has $x^2 < Cx$ if x is small enough; *how* small is actually small enough depends on C. The smaller the constant C, the closer we have to keep x to 0 to be sure

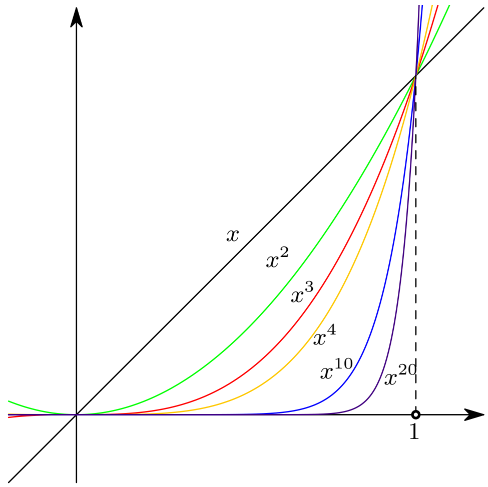


Figure 4. How the powers stack up. All graphs of $y = x^n$ ($n > 1$) are tangent to the x-axis at the origin. But the larger the exponent n the "flatter" the graph of $y = x^n$ is.

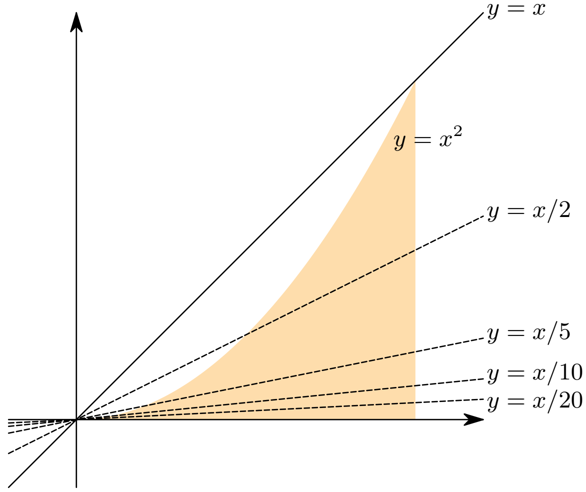


Figure 5. x^2 is smaller than any multiple of x, if x is small enough. Compare the quadratic function $y = x^2$ with a linear function $y = Cx$. Their graphs are a parabola and a straight line. Parts of the parabola may lie above the line, but as $x \searrow 0$ the parabola will always duck underneath the line.

that x^2 is smaller than Cx. Nevertheless, no matter how small C is, the parabola will eventually always reach the region below the line $y = Cx$.

8.6. Example: Little-oh arithmetic is a little funny. Both x^2 and x^3 are functions that are $o(x)$, i.e.

$$x^2 = o(x) \quad \text{and} \quad x^3 = o(x)$$

Nevertheless $x^2 \neq x^3$. So in working with little-oh we are giving up on the principle that says that two things that both equal a third object must themselves be equal; in other words, $a = b$ and $b = c$ implies $a = c$, but not when we're using little-ohs! We can also put it like this: just because two quantities both are much smaller than x, they don't have to be equal. In particular,

we can never cancel little-ohs!!!

In other words, the following is pretty wrong

$$o(x^2) - o(x^2) = 0.$$

Why? The two $o(x^2)$'s both refer to functions $h(x)$ that satisfy $\lim_{x\to 0} h(x)/x^2 = 0$, but there are many such functions, and the two $o(x^2)$'s could be abbreviations for different functions $h(x)$.

Contrast this with the following computation, which at first sight looks wrong even though it is actually right:

$$o(x^2) - o(x^2) = o(x^2).$$

In words: if we subtract two quantities both of which are negligible compared to x^2 for small x then the result will also be negligible compared to x^2 for small x.

8.7. Computations with Taylor polynomials. The following theorem is very useful because it lets us compute Taylor polynomials of a function without differentiating it.

8.8. Theorem. *If $f(x)$ and $g(x)$ are $n+1$ times differentiable functions then*

$$T_n f(x) = T_n g(x) \iff f(x) = g(x) + o(x^n). \tag{52}$$

In other words, if two functions have the same nth degree Taylor polynomial, then their difference is much smaller than x^n, at least, if x is small.

In principle the definition of $T_n f(x)$ lets us compute as many terms of the Taylor polynomial as we want, but in many (most) examples the computations quickly get out of hand. The following example shows what can happen.

8.9. How *NOT* to compute the Taylor polynomial of degree 12 of $f(x) = 1/(1+x^2)$. Diligently computing derivatives one by one we find

$$\begin{aligned}
f(x) &= \frac{1}{1+x^2} && \text{so } f(0) = 1\\
f'(x) &= \frac{-2x}{(1+x^2)^2} && \text{so } f'(0) = 0\\
f''(x) &= \frac{6x^2-2}{(1+x^2)^3} && \text{so } f''(0) = -2\\
f^{(3)}(x) &= 24\frac{x-x^3}{(1+x^2)^4} && \text{so } f^{(3)}(0) = 0\\
f^{(4)}(x) &= 24\frac{1-10x^2+5x^4}{(1+x^2)^5} && \text{so } f^{(4)}(0) = 24 = 4!\\
f^{(5)}(x) &= 240\frac{-3x+10x^3-3x^5}{(1+x^2)^6} && \text{so } f^{(4)}(0) = 0\\
f^{(6)}(x) &= -720\frac{-1+21x^2-35x^4+7x^6}{(1+x^2)^7} && \text{so } f^{(4)}(0) = -720 = -6!\\
&\ \ \vdots
\end{aligned}$$

Here we give up – can you find $f^{(12)}(x)$? After a lot of work all we have found is

$$T_6\left\{\frac{1}{1+x^2}\right\} = 1 - x^2 + x^4 - x^6.$$

By the way,

$$f^{(12)}(x) = 479001600\frac{1-78\,x^2+715\,x^4-1716\,x^6+1287\,x^8-286\,x^{10}+13\,x^{12}}{(1+x^2)^{13}}$$

and $479001600 = 12!$.

8.10. A much easier approach to finding the Taylor polynomial of any degree of $f(x) = 1/(1+x^2)$. Start with the Geometric Series: if $g(t) = 1/(1-t)$ then

$$g(t) = 1 + t + t^2 + t^3 + t^4 + \cdots + t^n + o(t^n).$$

Now substitute $t = -x^2$ in this limit,

$$g(-x^2) = 1 - x^2 + x^4 - x^6 + \cdots + (-1)^n x^{2n} + o\left(\left(-x^2\right)^n\right)$$

Since $o\left(\left(-x^2\right)^n\right) = o(x^{2n})$ and

$$g(-x^2) = \frac{1}{1-(-x^2)} = \frac{1}{1+x^2},$$

we have found

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \cdots + (-1)^n x^{2n} + o(x^{2n})$$

By Theorem (8.8) this implies

$$T_{2n}\left\{\frac{1}{1+x^2}\right\} = 1 - x^2 + x^4 - x^6 + \cdots + (-1)^n x^{2n}.$$

8.11. Example of multiplication of Taylor polynomials. Finding the Taylor polynomials of $e^{2x}/(1+x)$ directly from the definition is another recipe for headaches. Instead, we should exploit our knowledge of the Taylor polynomials of both factors e^{2x} and $1/(1+x)$:

$$\begin{aligned} e^{2x} &= 1 + 2x + \frac{2^2x^2}{2!} + \frac{2^3x^3}{3!} + \frac{2^4x^4}{4!} + o(x^4) \\ &= 1 + 2x + 2x^2 + \frac{4}{3}x^3 + \frac{2}{3}x^4 + o(x^4) \\ \frac{1}{1+x} &= 1 - x + x^2 - x^3 + x^4 + o(x^4). \end{aligned}$$

Then multiply these two

$$e^{2x}\cdot\frac{1}{1+x} = \left(1 + 2x + 2x^2 + \frac{4}{3}x^3 + \frac{2}{3}x^4 + o(x^4)\right)\cdot\left(1 - x + x^2 - x^3 + x^4 + o(x^4)\right)$$

$$\begin{array}{cccccccccccc} = & 1 & - & x & + & x^2 & - & x^3 & + & x^4 & + & o(x^4) \\ & & + & 2x & - & 2x^2 & + & 2x^3 & - & 2x^4 & + & o(x^4) \\ & & & & + & 2x^2 & - & 2x^3 & + & 2x^4 & + & o(x^4) \\ & & & & & & + & \frac{4}{3}x^3 & - & \frac{4}{3}x^4 & + & o(x^4) \\ & & & & & & & & + & \frac{2}{3}x^4 & + & o(x^4) \end{array}$$

$$= 1 + x + x^2 + \frac{1}{3}x^3 + \frac{1}{3}x^4 + o(x^4) \quad (x \to 0)$$

We conclude that

$$T_4\Big[\frac{e^{2x}}{1+x}\Big] = 1 + x + x^2 + \frac{1}{3}x^3 + \frac{1}{3}x^4.$$

8.12. Taylor's formula and Fibonacci numbers. The Fibonacci numbers are defined as follows: the first two are $f_0 = 1$ and $f_1 = 1$, and the others are defined by the equation

$$\boxed{f_n = f_{n-1} + f_{n-2}} \qquad \text{(Fib)}$$

So

$$\begin{aligned} f_2 &= f_1 + f_0 = 1 + 1 = 2,\\ f_3 &= f_2 + f_1 = 2 + 1 = 3,\\ f_4 &= f_3 + f_2 = 3 + 2 = 5,\\ &\text{etc.} \end{aligned}$$

The equation (Fib) lets us compute the whole sequence of numbers, one by one, when we are given only the first few numbers of the sequence (f_0 and f_1 in this case). Such an equation for the elements of a sequence is called a ***recursion relation***.

Now consider the function

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

Let

$$T_\infty f(x) = c_0 + c_1x + c_2x^2 + c_3x^3 + \cdots$$

be its Taylor series.

Due to Lagrange's remainder theorem we have, for any n,

$$\frac{1}{1-x-x^2} = c_0 + c_1x + c_2x^2 + c_3x^3 + \cdots + c_nx^n + o(x^n) \quad (x \to 0).$$

Multiply both sides with $1 - x - x^2$ and we get

$$\begin{aligned} 1 &= (1-x-x^2)\cdot(c_0 + c_1x + c_2x^2 + \cdots + c_nx^n + o(x^n)) \qquad (x\to 0)\\ &= \begin{array}{cccccccccccc} c_0 & + & c_1x & + & c_2x^2 & + & \cdots & + & c_nx^n & + & o(x^n) & \\ & - & c_0x & - & c_1x^2 & - & \cdots & - & c_{n-1}x^n & + & o(x^n) & \\ & & & - & c_0x^2 & - & \cdots & - & c_{n-2}x^n & - & o(x^n) & \qquad (x\to 0) \end{array}\\ &= c_0 + (c_1-c_0)x + (c_2-c_1-c_0)x^2 + (c_3-c_2-c_1)x^3 + \cdots\\ &\qquad \cdots + (c_n - c_{n-1} - c_{n-2})x^n + o(x^n) \quad (x\to 0) \end{aligned}$$

Compare the coefficients of powers x^k on both sides for $k = 0, 1, \ldots, n$ and we find

$$c_0 = 1, \quad c_1 - c_0 = 0 \implies c_1 = c_0 = 1, \quad c_2 - c_1 - c_0 = 0 \implies c_2 = c_1 + c_0 = 2$$

and in general

$$c_n - c_{n-1} - c_{n-2} = 0 \implies c_n = c_{n-1} + c_{n-2}$$

Therefore the coefficients of the Taylor series $T_\infty f(x)$ are exactly the Fibonacci numbers:

$$c_n = f_n \text{ for } n = 0, 1, 2, 3, \ldots$$

Since it is much easier to compute the Fibonacci numbers one by one than it is to compute the derivatives of $f(x) = 1/(1-x-x^2)$, this is a better way to compute the Taylor series of $f(x)$ than just directly from the definition.

8.13. More about the Fibonacci numbers. In this example we'll see a trick that lets us compute the Taylor series of ***any rational function***. You already know the trick: find the partial fraction decomposition of the given rational function. Ignoring the case that we have quadratic expressions in the denominator, this lets us represent our rational function as a sum of terms of the form

$$\frac{A}{(x-a)^p}.$$

These are easy to differentiate any number of times, and thus they allow us to write their Taylor series.

Let's apply this to the function $f(x) = 1/(1-x-x^2)$ from the example 8.12. First we factor the denominator.

$$1-x-x^2 = 0 \iff x^2+x-1=0 \iff x = \frac{-1\pm\sqrt{5}}{2}.$$

The number

$$\phi = \frac{1+\sqrt{5}}{2} \approx 1.618\,033\,988\,749\,89\ldots$$

is called the ***Golden Ratio***. It satisfies[2]

$$\phi + \frac{1}{\phi} = \sqrt{5}.$$

The roots of our polynomial x^2+x-1 are therefore

$$x_- = \frac{-1-\sqrt{5}}{2} = -\phi, \qquad x_+ = \frac{-1+\sqrt{5}}{2} = \frac{1}{\phi}.$$

and we can factor $1-x-x^2$ as follows

$$1-x-x^2 = -(x^2+x-1) = -(x-x_-)(x-x_+) = -(x-\frac{1}{\phi})(x+\phi).$$

So $f(x)$ can be written as

$$f(x) = \frac{1}{1-x-x^2} = \frac{-1}{(x-\frac{1}{\phi})(x+\phi)} = \frac{A}{x-\frac{1}{\phi}} + \frac{B}{x+\phi}$$

The Heaviside trick will tell us what A and B are, namely,

$$A = \frac{-1}{\frac{1}{\phi}+\phi} = \frac{-1}{\sqrt{5}}, \qquad B = \frac{1}{\frac{1}{\phi}+\phi} = \frac{1}{\sqrt{5}}$$

The nth derivative of $f(x)$ is

$$f^{(n)}(x) = \frac{A(-1)^n n!}{\left(x-\frac{1}{\phi}\right)^{n+1}} + \frac{B(-1)^n n!}{(x+\phi)^{n+1}}$$

Setting $x=0$ and dividing by $n!$ finally gives us the coefficient of x^n in the Taylor series of $f(x)$. The result is the following formula for the nth Fibonacci number

$$c_n = \frac{f^{(n)}(0)}{n!} = \frac{1}{n!}\frac{A(-1)^n n!}{\left(-\frac{1}{\phi}\right)^{n+1}} + \frac{1}{n!}\frac{B(-1)^n n!}{(\phi)^{n+1}} = -A\phi^{n+1} - B\left(-\frac{1}{\phi}\right)^{n+1}$$

---

[2] To prove this, use $\frac{1}{\phi} = \frac{2}{1+\sqrt{5}} = \frac{2}{1+\sqrt{5}}\frac{1-\sqrt{5}}{1-\sqrt{5}} = \frac{-1+\sqrt{5}}{2}$.

Using the values for A and B we find

$$f_n = c_n = \frac{1}{\sqrt{5}}\left\{\phi^{n+1} - (-\frac{1}{\phi})^{n+1}\right\} \tag{53}$$

9. Problems

Are the following statements True or False? In mathematics this means that we should either show that the statement always holds or else give at least one counterexample, thereby showing that the statement is not always true.

1. $(1+x^2)^2 - 1 = o(x)$?

2. $(1+x^2)^2 - 1 = o(x^2)$?

3. $\sqrt{1+x} - \sqrt{1-x} = o(x)$?

4. $o(x) + o(x) = o(x)$?

5. $o(x) - o(x) = o(x)$?

6. $o(x) \cdot o(x) = o(x)$?

7. $o(x^2) + o(x) = o(x^2)$?

8. $o(x^2) - o(x^2) = o(x^3)$?

9. $o(2x) = o(x)$?

10. $o(x) + o(x^2) = o(x)$?

11. $o(x) + o(x^2) = o(x^2)$?

12. $1 - \cos x = o(x)$?

13. Define

$$f(x) = \begin{cases} e^{-1/x^2} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

This function goes to zero *very* quickly as $x \to 0$ but is 0 only at 0. Prove that $f(x) = o(x^n)$ for every n.

14. For which value(s) of k are the following true (as $x \to 0$)?

(a) $\sqrt{1+x^2} = 1 + o(x^k)$

(b) $\sqrt[3]{1+x^2} = 1 + o(x^k)$

(c) $1 - \cos(x^2) = o(x^k)$

(d) $1 - \big(\cos x\big)^2 = o(x^k)$

15. **[Group Problem]** Let g_n be the coefficient of x^n in the Taylor series of the function

$$g(x) = \frac{1}{2 - 3x + x^2}$$

(a) Compute g_0 and g_1 directly from the definition of the Taylor series.

(b) Show that the recursion relation $g_n = 3g_{n-1} - 2g_{n-2}$ holds for all $n \geq 2$.

(c) Compute g_2, g_3, g_4, g_5.

(d) Using a partial fraction decomposition of $g(x)$ find a formula for $g^{(n)}(0)$, and hence for g_n.

16. Answer the same questions as in the previous problem, for the functions

$$h(x) = \frac{x}{2 - 3x + x^2}$$

and

$$k(x) = \frac{2 - x}{2 - 3x + x^2}.$$

17. Let h_n be the coefficient of x^n in the Taylor series of

$$h(x) = \frac{1 + x}{2 - 5x + 2x^2}.$$

(a) Find a recursion relation for the h_n.

(b) Compute $h_0, h_1, \ldots, h_8$.

(c) Derive a formula for h_n valid for all n, by using a partial fraction expansion.

(d) Is h_{2009} more or less than a million? A billion?

Find the Taylor series for the following functions, by substituting, adding, multiplying, applying long division and/or differentiating known series for $\frac{1}{1+x}$, e^x, $\sin x$, $\cos x$ and $\ln x$.

18. e^{at}

19. e^{1+t}

20. e^{-t^2}

21. $\dfrac{1+t}{1-t}$

22. $\dfrac{1}{1+2t}$

23.

$$f(x) = \begin{cases} \frac{\sin(x)}{x} & \text{if } x \neq 0 \\ 1 & \text{if } x = 0 \end{cases}$$

24. $\dfrac{\ln(1+x)}{x}$

25. $\dfrac{e^t}{1-t}$

26. $\dfrac{1}{\sqrt{1-t}}$

27. $\dfrac{1}{\sqrt{1-t^2}}$ (recommendation: use the answer to problem 9.26)

28. $\arcsin t$
(use problem 9.26 again)

29. Compute $T_4[e^{-t}\cos t]$ (See example 8.11.)

30. $T_4[e^{-t}\sin 2t]$

31. $\dfrac{1}{2-t-t^2}$

32. $\sqrt[3]{1+2t+t^2}$

33. $\ln(1-t^2)$

34. $\sin t\cos t$

10. Differentiating and Integrating Taylor polynomials

If

$$T_n f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n$$

is the Taylor polynomial of a function $y = f(x)$, then what is the Taylor polynomial of its derivative $f'(x)$?

10.1. Theorem. *The Taylor polynomial of degree $n-1$ of $f'(x)$ is given by*

$$T_{n-1}\{f'(x)\} = a_1 + 2a_2x + \cdots + na_nx^{n-1}.$$

In other words, "the Taylor polynomial of the derivative is the derivative of the Taylor polynomial."

Written in terms of little-oh notation the theorem says that if f is an n times differentiable function

$$f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n + o(x^n)$$
$$\implies f'(x) = a_1 + 2a_2x + \cdots + na_nx^{n-1} + o(x^{n-1}).$$

Proof. Let $g(x) = f'(x)$. Then $g^{(k)}(0) = f^{(k+1)}(0)$, so that

$$T_{n-1}g(x) = g(0) + g'(0)x + g^{(2)}(0)\frac{x^2}{2!} + \cdots + g^{(n-1)}(0)\frac{x^{n-1}}{(n-1)!}$$

$$= f'(0) + f^{(2)}(0)x + f^{(3)}(0)\frac{x^2}{2!} + \cdots + f^{(n)}(0)\frac{x^{n-1}}{(n-1)!} \qquad (\$)$$

On the other hand, if $T_nf(x) = a_0 + a_1x + \cdots + a_nx^n$, then $a_k = f^{(k)}(0)/k!$, so that

$$ka_k = \frac{k}{k!}f^{(k)}(0) = \frac{f^{(k)}(0)}{(k-1)!}.$$

In other words,

$$1\cdot a_1 = f'(0),\ 2a_2 = f^{(2)}(0),\ 3a_3 = \frac{f^{(3)}(0)}{2!},\ \text{etc.}$$

So, continuing from ($) you find that

$$T_{n-1}\{f'(x)\} = T_{n-1}g(x) = a_1 + 2a_2x + \cdots + na_nx^{n-1}$$

as claimed. □

10.2. Example: Taylor polynomial of $(1-x)^{-2}$. We compute the Taylor polynomial of $f(x) = 1/(1-x)^2$ by noting that

$$f(x) = F'(x), \text{ where } F(x) = \frac{1}{1-x}.$$

Since

$$T_{n+1}F(x) = 1 + x + x^2 + x^3 + \cdots + x^{n+1},$$

theorem 10.1 implies that

$$T_n\left\{\frac{1}{(1-x)^2}\right\} = 1 + 2x + 3x^2 + 4x^3 + \cdots + (n+1)x^n$$

10.3. Example: Taylor polynomials of $\arctan x$. Since integration undoes differentiation we can use Theorem 10.1 to find the Taylor polynomials of integrals of functions whose Taylor polynomials we already know. The following example shows that there is one difference, namely, when we integrate a Taylor expansion we need to determine the integration constant.

Let $f(x) = \arctan x$. Then we know that

$$f'(x) = \frac{1}{1+x^2}.$$

By substitution of $t = -x^2$ in the Taylor polynomial of $1/(1-t)$ we had found

$$T_{2n}\{f'(x)\} = T_{2n}\left\{\frac{1}{1+x^2}\right\} = 1 - x^2 + x^4 - x^6 + \cdots + (-1)^n x^{2n}.$$

This Taylor polynomial must be the derivative of $T_{2n+1}f(x)$, so we have

$$T_{2n+1}\{\arctan x\} = C + x - \frac{x^3}{3} + \frac{x^5}{5} + \cdots + (-1)^n\frac{x^{2n+1}}{2n+1},$$

where C is the integration constant. Normally we would write the integration constant at the end, but here we have written it at the beginning of the sum. This makes no difference to the sum of course, but it makes sense to do this because C is a constant, and in Taylor polynomials we habitually write the constant term first. This shows us that to find C we merely have to set x equal to zero:

$$C = \arctan(0) = 0.$$

Therefore we get

$$T_{2n+1}\{\arctan x\} = x - \frac{x^3}{3} + \frac{x^5}{5} + \cdots + (-1)^n\frac{x^{2n+1}}{2n+1}.$$

11. Problems on Integrals and Taylor Expansions

1. (a) Compute $T_2\{\sin t\}$ and give an upper bound for $R_2\{\sin t\}$ for $0 \le t \le 0.5$

(b) Use part **(a)** to approximate

$$\int_0^{0.5} \sin(x^2)\,dx,$$

and give an upper bound for the error in your approximation.

2. (a) Find the second degree Taylor polynomial for the function e^t.

(b) Use it to give an estimate for the integral

$$\int_0^1 e^{x^2}\,dx$$

(c) Suppose instead we used the 5th degree Taylor polynomial $p(t) = T_5\{e^t\}$ for e^t to give an estimate for the integral:

$$\int_0^1 e^{x^2}\,dx$$

Give an upper bound for the error:

$$\left|\int_0^1 e^{x^2}\,dx - \int_0^1 p(x^2)\,dx\right|$$

Note: You need not find $p(t)$ or the integral $\int_0^1 p(x^2)\,dx$.

3. Approximate $\displaystyle\int_0^{0.1} \arctan x\,dx$ and estimate the error in your approximation by analyzing $T_2 f(t)$ and $R_2 f(t)$ where $f(t) = \arctan t$.

4. Approximate $\displaystyle\int_0^{0.1} x^2 e^{-x^2}\,dx$ and estimate the error in your approximation by analyzing $T_3 f(t)$ and $R_3 f(t)$ where $f(t) = te^{-t}$.

5. Estimate $\displaystyle\int_0^{0.5} \sqrt{1+x^4}\,dx$ with an error of less than 10^{-4}.

6. Estimate $\displaystyle\int_0^{0.1} \arctan x\,dx$ with an error of less than 0.001.

12. Proof of Theorem 8.8

12.1. Lemma. *If $h(x)$ is a k times differentiable function on some interval containing 0, and if for some integer $k < n$ one has $h(0) = h'(0) = \cdots = h^{(k-1)}(0) = 0$, then*

$$\lim_{x\to 0} \frac{h(x)}{x^k} = \frac{h^{(k)}(0)}{k!}. \tag{54}$$

Proof. Just apply l'Hopital's rule k times. You get

$$\lim_{x\to 0}\frac{h(x)}{x^k} \overset{=\frac{0}{0}}{=} \lim_{x\to 0}\frac{h'(x)}{kx^{k-1}} \overset{=\frac{0}{0}}{=} \lim_{x\to 0}\frac{h^{(2)}(x)}{k(k-1)x^{k-2}} \overset{=\frac{0}{0}}{=} \cdots$$

$$\cdots = \lim_{x\to 0}\frac{h^{(k-1)}(x)}{k(k-1)\cdots 2x^1} \overset{=\frac{0}{0}}{=} \frac{h^{(k)}(0)}{k(k-1)\cdots 2\cdot 1}$$

□

First define the function $h(x) = f(x) - g(x)$. If $f(x)$ and $g(x)$ are n times differentiable, then so is $h(x)$.

The condition $T_n f(x) = T_n g(x)$ means that

$$f(0) = g(0), \quad f'(0) = g'(0), \quad \ldots, \quad f^{(n)}(0) = g^{(n)}(0),$$

which says, in terms of $h(x)$,

$$h(0) = h'(0) = h''(0) = \cdots = h^{(n)}(0) = 0, \tag{†}$$

i.e.

$$T_n h(x) = 0.$$

We now prove the first part of the theorem: suppose $f(x)$ and $g(x)$ have the same nth degree Taylor polynomial. Then we have just argued that $T_n h(x) = 0$, and Lemma 12.1 (with $k = n$) says that $\lim_{x\to 0} h(x)/x^n = 0$, as claimed.

To conclude we show the converse also holds. So suppose that $\lim_{x\to 0} h(x)/x^n = 0$. We'll show that (†) follows. If (†) were not true then there would be a smallest integer $k \le n$ such that

$$h(0) = h'(0) = h''(0) = \cdots = h^{(k-1)}(0) = 0, \text{ but } h^{(k)}(0) \ne 0.$$

This runs into the following contradiction with Lemma 12.1

$$0 \neq \frac{h^{(k)}(0)}{k!} = \lim_{x\to 0}\frac{h(x)}{x^k} = \lim_{x\to 0}\frac{h(x)}{x^n}\cdot\frac{x^n}{x^k} = 0\cdot\underbrace{\lim_{x\to 0} x^{n-k}}_{(*)} = 0.$$

Here the limit $(*)$ exists because $n \geq k$.

13. Proof of Lagrange's formula for the remainder

For simplicity assume $x > 0$. Consider the function

$$g(t) = f(0) + f'(0)t + \frac{f''(0)}{2}t^2 + \cdots + \frac{f^{(n)}(0)}{n!}t^n + Kt^{n+1} - f(t),$$

where

$$K \stackrel{\text{def}}{=} -\frac{f(0) + f'(0)x + \cdots + \frac{f^{(n)}(0)}{n!}x^n - f(x)}{x^{n+1}} \tag{55}$$

We have chosen this particular K to be sure that

$$g(x) = 0.$$

Just by computing the derivatives we also find that

$$g(0) = g'(0) = g''(0) = \cdots = g^{(n)}(0) = 0,$$

while

$$g^{(n+1)}(t) = (n+1)!K - f^{(n+1)}(t). \tag{56}$$

We now apply *Rolle's Theorem* n times:

- since $g(t)$ vanishes at $t = 0$ and at $t = x$ there exists an x_1 with $0 < x_1 < x$ such that $g'(x_1) = 0$
- since $g'(t)$ vanishes at $t = 0$ and at $t = x_1$ there exists an x_2 with $0 < x_2 < x_1$ such that $g'(x_2) = 0$
- since $g''(t)$ vanishes at $t = 0$ and at $t = x_2$ there exists an x_3 with $0 < x_3 < x_2$ such that $g''(x_3) = 0$

  $\vdots$
- since $g^{(n)}(t)$ vanishes at $t = 0$ and at $t = x_n$ there exists an x_{n+1} with $0 < x_{n+1} < x_n$ such that $g^{(n)}(x_{n+1}) = 0$.

We now set $\xi = x_{n+1}$, and observe that we have shown that $g^{(n+1)}(\xi) = 0$, so by (56) we get

$$K = \frac{f^{(n+1)}(\xi)}{(n+1)!}.$$

Apply that to (55) and we finally get

$$f(x) = f(0) + f'(0)x + \cdots + \frac{f^{(n)}(0)}{n!}x^n + \frac{f^{(n+1)}(\xi)}{(n+1)!}x^{n+1}.$$

CHAPTER 5

Sequences and Series

1. Introduction

1.1. A different point of view on Taylor expansions. In the previous chapter we saw that certain functions, like e^x can be approximated by computing their Taylor expansions,

$$\begin{aligned}
e^x &= 1 + x + R_1(x)\\
e^x &= 1 + x + \frac{x^2}{2!} + R_2(x)\\
e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + R_3(x)\\
&\ \vdots\\
e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + R_n(x).
\end{aligned}$$

We found a formula for the "remainder" or "error in the approximation" $R_n(x)$, and for fixed n we considered the rate at which this error vanishes when $x \to 0$. In this chapter we consider a fixed value of x and let $n \to \infty$. If the remainder term R_n gets smaller and eventually goes to zero when $n \to \infty$, then we could say that

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + \cdots \tag{57}$$

where the $\cdots$ indicate that we are adding ***infinitely many terms***. For instance, if $x = 1$ then the above formula would say that

$$e = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \cdots + \frac{1}{n!} + \cdots \tag{58}$$

In other words the number e can be written as the sum of infinitely many fractions. There are many other formulas like this, e.g. Leibniz' formula for $\pi/4$,

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \cdots \tag{59}$$

which we get by setting $x = 1$ in the Taylor expansion for $\arctan x$.

Such sums with infinitely many terms are called "series" and it turns out their applications go far beyond the pretty formulas for famous numbers like (58) or (59).

1.2. Some sums with infinitely many terms. Before we start using infinite sums we should take a good look at what it means to "add infinitely many numbers." It is not at all clear that this concept is well defined. For example, the sum

$$1 + 1 + 1 + 1 + \cdots$$

clearly becomes larger and larger as we keep on adding more terms. Adding infinitely many ones together should give us an infinitely large sum. So we might like to say that

$$1+1+1+1+\cdots=\infty.$$

Since we do not know what ∞ is ("∞ is not a number"), the sum $1+1+1+\cdots$ is not defined, or "does not exist."

This may seem straightforward and it looks like this is the only possible way to consider the sum $1+1+\cdots$. The following sum is trickier:

$$1-1+1-1+1-1+1-1+1\cdots$$

If we group the terms like this

$$(1-1)+(1-1)+(1-1)+(1-1)+\cdots$$

then $1-1=0$ tells us that the sum should be

$$(1-1)+(1-1)+(1-1)+(1-1)+\cdots=0+0+0+\cdots=0. \tag{60}$$

On the other hand, if we group the terms according to

$$1\ \underbrace{-1+1}_{=0}\ \underbrace{-1+1}_{=0}\ \underbrace{-1+1}_{=0}-\cdots$$

and we get

$$1\ \underbrace{-1+1}\ \underbrace{-1+1}\ \underbrace{-1+1}-\cdots=1+0+0+0+\cdots=1. \tag{61}$$

Look carefully at (60) and (61): depending on how we group the terms the sum can be either 0 or 1! Clearly the infinite sum $1-1+1-1+1-\cdots$ does not make a lot of sense.

Apparently things that we always do with finite sums, like rearranging the terms, or changing the order in which we add the terms, do not work for infinite sums.

We are left with the question *which infinite sums can we add?* Our approach will be to think of an infinite sum as the limit of finite sums. For example, to add the sum (58) we consider the results of adding the first few terms:

$$\begin{aligned}
1&=1\\
1+\frac{1}{1!}&=2\\
1+\frac{1}{1!}+\frac{1}{2!}&=2.5\\
1+\frac{1}{1!}+\frac{1}{2!}+\frac{1}{3!}&=2.666\ldots\\
1+\frac{1}{1!}+\frac{1}{2!}+\frac{1}{3!}+\frac{1}{4!}&=2.7083333\ldots\\
1+\frac{1}{1!}+\frac{1}{2!}+\frac{1}{3!}+\frac{1}{4!}+\frac{1}{5!}&=2.716666\ldots
\end{aligned}$$

We see that as we include more terms in the sum its value appears to get closer to a specific value. This is how we will interpret "the sum of infinitely many numbers": the whole sum is the limit for $n\to\infty$ of what we get by adding the first n terms. More precisely, if we write s_n for the sum of the first n terms in an infinite sum, then we will define the entire sum to be $\lim_{n\to\infty} s_n$. The next two sections go into the details of these definitions. First we discuss "the limit for $n\to\infty$" of certain quantities like s_n (which we will call "sequences"). The discussion is similar to that of limits of functions $\lim_{x\to a} f(x)$, except that now the variable x is replaced by a variable n that only takes integer values.

Once we have a better understanding of limits of sequences, we return to infinite sums in Section 4.

2. Sequences

We shall call a ***sequence*** any ordered sequence of numbers $a_1, a_2, a_3, \ldots$. For each positive integer n we have to specify a number a_n. It sometimes helps to think of a sequence as a function $a = a(n)$ whose domain consists only of the integers $\{1, 2, 3, 4, \ldots\}$, and for which we use subscript notation a_n instead of the more common notation $a(n)$ for functions.

2.1. Examples of sequences.

definition	first few numbers in the sequence
$a_n = n$	$1, 2, 3, 4, \ldots$
$b_n = 0$	$0, 0, 0, 0, \ldots$
$c_n = \frac{1}{n}$	$\frac{1}{1}, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$
$d_n = \left(-\frac{1}{3}\right)^n$	$-\frac{1}{3}, \frac{1}{9}, -\frac{1}{27}, \frac{1}{81}, \ldots$
$E_n = 1 + \dfrac{1}{1!} + \dfrac{1}{2!} + \dfrac{1}{3!} + \cdots + \dfrac{1}{n!}$	$1, 2, 2\frac{1}{2}, 2\frac{2}{3}, 2\frac{17}{24}, 2\frac{43}{60}, \ldots$
$S_n = x - \dfrac{x^3}{3!} + \cdots + (-1)^n \dfrac{x^{2n+1}}{(2n+1)!}$	$x, x - \dfrac{x^3}{3!}, x - \dfrac{x^3}{3!} + \dfrac{x^5}{5!}, \ldots$

The last two sequences are derived from the Taylor polynomials of e^x (at $x = 1$) and $\sin x$ (at any x). The last example S_n really is a sequence of functions, i.e. the n^{th} number in the sequence depends on x.

2.2. Definition. *A sequence of numbers $(a_n)_{n=1}^{\infty}$ converges to a limit L, if for every $\epsilon > 0$ there is a number N_ϵ such that for all $n > N_\epsilon$ one has*

$$|a_n - L| < \epsilon.$$

One writes

$$\lim_{n\to\infty} a_n = L$$

2.3. Example: $\lim_{n\to\infty} \dfrac{1}{n} = 0$. The sequence $c_n = 1/n$ converges to 0. To prove this let $\epsilon > 0$ be given. We have to find an N_ϵ such that

$$|c_n| < \epsilon \text{ for all } n > N_\epsilon.$$

The c_n are all positive, so $|c_n| = c_n$, and hence

$$|c_n| < \epsilon \iff \frac{1}{n} < \epsilon \iff n > \frac{1}{\epsilon},$$

which prompts us to choose $N_\epsilon = 1/\epsilon$. The calculation we just did shows that if $n > \frac{1}{\epsilon} = N_\epsilon$, then $|c_n| < \epsilon$. That means that $\lim_{n\to\infty} c_n = 0$.

2.4. Example: $\lim_{n\to\infty} a^n = 0$ if $|a| < 1$. As in the previous example one can show that $\lim_{n\to\infty} 2^{-n} = 0$, and more generally, that for any constant a with $-1 < a < 1$ one has

$$\lim_{n\to\infty} a^n = 0.$$

Indeed,

$$|a^n| = |a|^n = e^{n \ln |a|} < \epsilon$$

holds if and only if

$$n \ln |a| < \ln \epsilon.$$

Since $|a| < 1$ we have $\ln |a| < 0$ so that dividing by $\ln |a|$ reverses the inequality, with result

$$|a^n| < \epsilon \iff n > \frac{\ln \epsilon}{\ln |a|}$$

The choice $N_\epsilon = (\ln \epsilon)/(\ln |a|)$ therefore guarantees that $|a^n| < \epsilon$ whenever $n > N_\epsilon$.

The case $|a| \geq 1$ (without proof). If $a > 1$ then the quantity a^n grows larger with increasing n, and the limit $\lim_{n\to\infty} a^n$ does not exist.

When $a \leq -1$ then the sequence of numbers $1, a, a^2, a^3, \ldots$ flip-flops between positive and negative numbers, while the absolute value $|a^n| = |a|^n$ either becomes infinitely large (when $a < -1$), or else remains exactly equal to 1 (when $a = -1$). In either case the limit $\lim_{n\to\infty} a^n$ does not exist.

Finally, if $a = +1$, then the sequence $1, a, a^2, \ldots$ is very simple, namely, $a^n = 1$ for all n. Clearly in this case $\lim_{n\to\infty} a^n = \lim_{n\to\infty} 1 = 1$.

One can show that the operation of taking limits of sequences obeys the same rules as taking limits of functions.

2.5. Theorem. *If*

$$\lim_{n\to\infty} a_n = A \text{ and } \lim_{n\to\infty} b_n = B,$$

then one has

$$\begin{aligned}
\lim_{n\to\infty} a_n \pm b_n &= A \pm B \\
\lim_{n\to\infty} a_n b_n &= AB \\
\lim_{n\to\infty} \frac{a_n}{b_n} &= \frac{A}{B} \qquad \text{(assuming } B \neq 0\text{)}.
\end{aligned}$$

The so-called "sandwich theorem" for ordinary limits also applies to limits of sequences. Namely, one has

2.6. Sandwich theorem. *If a_n is a sequence which satisfies $b_n < a_n < c_n$ for all n, and if $\lim_{n\to\infty} b_n = \lim_{n\to\infty} c_n = 0$, then $\lim_{n\to\infty} a_n = 0$.*

Finally, one can show this:

2.7. Theorem. *If $f(x)$ is a function which is continuous at $x = A$, and a_n is a sequence which converges to A, then*

$$\lim_{n\to\infty} f(a_n) = f\left(\lim_{n\to\infty} a_n\right) = f(A).$$

2.8. Example. Since $\lim_{n\to\infty} 1/n = 0$ and since $f(x) = \cos x$ is continuous at $x = 0$ we have

$$\lim_{n\to\infty} \cos\frac{1}{n} = \cos\Big(\lim_{n\to\infty}\frac{1}{n}\Big) = \cos 0 = 1.$$

2.9. Limits of rational functions. You can compute the limit of any rational function of n by dividing numerator and denominator by the highest occurring power of n. Here is an example:

$$\lim_{n\to\infty}\frac{2n^2-1}{n^2+3n} = \lim_{n\to\infty}\frac{2-\left(\frac{1}{n}\right)^2}{1+3\cdot\frac{1}{n}} = \frac{2-0^2}{1+3\cdot 0^2} = 2.$$

2.10. Example. Application of the Sandwich theorem. We show that $\lim_{n\to\infty}\frac{1}{\sqrt{n^2+1}} = 0$ in two different ways.

Method 1: Since $\sqrt{n^2+1} > \sqrt{n^2} = n$ we have

$$0 < \frac{1}{\sqrt{n^2+1}} < \frac{1}{n}.$$

The sequences "0" and $\frac{1}{n}$ both go to zero, so the Sandwich theorem implies that $1/\sqrt{n^2+1}$ also goes to zero.

Method 2: Divide numerator and denominator both by n to get

$$a_n = \frac{1/n}{\sqrt{1+(1/n)^2}} = f\left(\frac{1}{n}\right), \qquad \text{where } f(x) = \frac{x}{\sqrt{1+x^2}}.$$

Since $f(x)$ is continuous at $x = 0$, and since $\frac{1}{n} \to 0$ as $n \to \infty$, we conclude that a_n converges to 0. You could write the computation like this:

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} f\left(\tfrac{1}{n}\right) = f\Big(\lim_{n\to\infty}\tfrac{1}{n}\Big) = f(0) = 0 \qquad \left(\begin{array}{c}\text{we would have to}\\ \text{say what } f \text{ is}\end{array}\right)$$

or like this:

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty}\frac{1/n}{\sqrt{1+(1/n)^2}} = \frac{\lim_{n\to\infty} 1/n}{\sqrt{\lim_{n\to\infty} 1+(1/n)^2}} = \frac{0}{\sqrt{1+0^2}} = 0.$$

2.11. Example: factorial beats any exponential. Factorials show up in Taylor's formula, so it is useful to know that $n!$ goes to infinity much faster than 2^n or 3^n or 100^n, or any x^n. In this example we'll show that

$$\lim_{n\to\infty}\frac{x^n}{n!} = 0 \text{ for any real number } x. \tag{62}$$

If $|x| \le 1$ then this is easy, for we would have $|x^n| \le 1$ for all $n \ge 0$ and thus

$$\left|\frac{x^n}{n!}\right| \le \frac{1}{n!} = \frac{1}{1\cdot\underbrace{2\cdot 3\cdots(n-1)\cdot n}_{n-1 \text{ factors}}} \le \frac{1}{1\cdot\underbrace{2\cdot 2\cdots 2\cdot 2}_{n-1\text{ factors}}} = \frac{1}{2^{n-1}}.$$

Written without the absolute values this says

$$-\left(\frac{1}{2}\right)^{n-1} \le \frac{x^n}{n!} \le \left(\frac{1}{2}\right)^{n-1}.$$

Both $(1/2)^{n-1}$ and $-(1/2)^{n-1}$ go to 0 as $n \to \infty$, so the Sandwich Theorem applies and tells us that (62) is true, at least when $|x| \le 1$.

If $|x| > 1$ then we need a slightly longer argument. For arbitrary x we first choose an integer $N \geq 2x$. Then for all $n \geq N$ one has

$$\begin{aligned} \frac{x^n}{n!} &\leq \frac{|x| \cdot |x| \cdots |x| \cdot |x|}{1 \cdot 2 \cdot 3 \cdots n} && \text{use } |x| \leq \frac{N}{2} \\ &\leq \frac{N \cdot N \cdot N \cdots N \cdot N}{1 \cdot 2 \cdot 3 \cdots n} \left(\frac{1}{2}\right)^n \end{aligned}$$

Split fraction into two parts, one containing the first N factors from both numerator and denominator, the other the remaining factors:

$$\underbrace{\frac{N}{1} \cdot \frac{N}{2} \cdot \frac{N}{3} \cdots \frac{N}{N}}_{=N^N/N!} \cdot \frac{N}{N+1} \cdots \frac{N}{n} = \frac{N^N}{N!} \cdot \underbrace{\frac{N}{N+1}}_{<1} \cdot \underbrace{\frac{N}{N+2}}_{<1} \cdots \underbrace{\frac{N}{n}}_{<1} \leq \frac{N^N}{N!}$$

Hence we have

$$\left|\frac{x^n}{n!}\right| \leq \frac{N^N}{N!} \left(\frac{1}{2}\right)^n$$

if $2|x| \leq N$ and $n \geq N$.

Here everything is independent of n, except for the last factor $(\frac{1}{2})^n$ which causes the whole thing to converge to zero as $n \to \infty$.

3. Problems on Limits of Sequences

Compute the following limits:

1. $\displaystyle\lim_{n\to\infty} \frac{n}{2n-3}$ •

2. $\displaystyle\lim_{n\to\infty} \frac{n^2}{2n-3}$ •

3. $\displaystyle\lim_{n\to\infty} \frac{n^2}{2n^2+n-3}$ •

4. $\displaystyle\lim_{n\to\infty} \frac{2^n+1}{1-2^n}$ •

5. $\displaystyle\lim_{n\to\infty} \frac{2^n+1}{1-3^n}$ •

6. $\displaystyle\lim_{n\to\infty} \frac{e^n+1}{1-2^n}$ •

7. $\displaystyle\lim_{n\to\infty} \frac{n^2}{(1.01)^n}$ •

8. $\displaystyle\lim_{n\to\infty} \frac{1000^n}{n!}$ •

9. $\displaystyle\lim_{n\to\infty} \frac{n!+1}{(n+1)!}$ •

10. **[Group Problem]** Compute $\displaystyle\lim_{n\to\infty} \frac{(n!)^2}{(2n)!}$
[Hint: write out all the factors in numerator and denominator.]

11. **[Group Problem]** Let f_n be the nth Fibonacci number. Compute

$$\lim_{n\to\infty} \frac{f_n}{f_{n-1}}$$

•

4. Series

4.1. Definitions.

A ***series*** is an infinite sum:

$$a_1 + a_2 + a_3 + \cdots = \sum_{k=1}^{\infty} a_k.$$

The numbers $a_1, a_2, \ldots$ are the terms of the sum. The result of adding the first n terms,

$$s_n = a_1 + a_2 + \cdots + a_n$$

is called the n^{th} ***partial sum***. The partial sums form themselves a sequence: $s_1, s_2, s_3, \ldots$ obtained by adding one extra term each time. Our plan, formulated in the beginning of this chapter, was to say that we have added the entire series provided we can take the limit of the partial sums s_n. Thus, by definition, we say *the series converges to a number S, which we call the **sum of the series,** if* $\lim_{n\to\infty} s_n$ *exists, and if*

$$S = \lim_{n\to\infty} s_n$$

i.e. if

$$S = \lim_{n\to\infty} a_1 + a_2 + \cdots + a_n.$$

If the limit does not exist, then we say the series ***diverges***. If the limit does exist, then we write either

$$S = a_1 + a_2 + a_3 + \cdots$$

or

$$S = \sum_{k=1}^{\infty} a_k.$$

4.2. Examples. It is difficult to find precise formulas for the partial sums s_n that different series can produce, and this limits the number of series that we can add "from scratch." In this section we present the few examples where some algebraic trick allows us to simplify the sum that defines s_n and then to take the limit of s_n as $n \to \infty$.

The geometric series. The geometric sum formula is one formula that lets us add partial sums of a particular series. For example, it tells us that

$$1 + \left(\tfrac{1}{2}\right) + \left(\tfrac{1}{2}\right)^2 + \left(\tfrac{1}{2}\right)^3 + \cdots + \left(\tfrac{1}{2}\right)^n = \frac{1 - \left(\frac{1}{2}\right)^{n+1}}{1 - \frac{1}{2}} = 2\Big(1 - \frac{1}{2^{n+1}}\Big) = 2 - \frac{1}{2^n}.$$

Since $\lim_{n\to\infty} \frac{1}{2^n} = 0$, we find

$$1 + \tfrac{1}{2} + \tfrac{1}{4} + \tfrac{1}{8} + \tfrac{1}{16} + \cdots = 2.$$

Telescoping series. It turns out that

$$\frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \frac{1}{3\cdot 4} + \cdots + \frac{1}{n(n+1)} + \cdots = 1.$$

There is a trick that allows us to compute the n^{th} partial sum. The trick begins with the miraculous observation that

$$\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}.$$

This allows us to rewrite the n^{th} partial sum as

$$\begin{aligned} s_n &= \frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \frac{1}{3\cdot 4} + \cdots + \frac{1}{n(n+1)} \\ &= \big(\frac{1}{1} - \frac{1}{2}\big) + \big(\frac{1}{2} - \frac{1}{3}\big) + \big(\frac{1}{3} - \frac{1}{4}\big) + \cdots + \big(\frac{1}{n} - \frac{1}{n+1}\big) \\ &= \frac{1}{1} \underbrace{-\frac{1}{2} + \frac{1}{2}}_{=0} \underbrace{-\frac{1}{3} + \frac{1}{3}}_{=0} + \cdots \underbrace{-\frac{1}{n} + \frac{1}{n}}_{=0} - \frac{1}{n+1} \\ &= 1 - \frac{1}{n+1}. \end{aligned}$$

This sum is called "telescoping" because in the last step of the computation almost all terms cancel and the whole sum collapses like a telescope to just two terms.

Once we have the formula for s_n it is easy to compute the sum:

$$\begin{aligned}
S &= \frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \frac{1}{3\cdot 4} + \cdots \\
&= \lim_{n\to\infty} \frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \frac{1}{3\cdot 4} + \cdots + \frac{1}{n(n+1)} \\
&= \lim_{n\to\infty} 1 - \frac{1}{n+1} \\
&= 1.
\end{aligned}$$

4.3. Properties of series. Just with limits, derivatives, integrals and sequences there are a number of properties that make it easier to work with series.

4.4. Theorem. *If the two series*

$$A = \sum_{k=1}^{\infty} a_k = a_1 + a_2 + a_3 + \cdots \quad \textit{and} \quad B = \sum_{k=1}^{\infty} b_k = b_1 + b_2 + b_3 + \cdots$$

both converge, then so does the series

$$\sum_{k=1}^{\infty} (a_k + b_k) = (a_1 + b_1) + (a_2 + b_2) + (a_3 + b_3) + \cdots .$$

Moreover, one has

$$\sum_{k=1}^{\infty} (a_k + b_k) = A + B, \qquad \textit{i.e.} \quad \sum_{k=1}^{\infty} (a_k + b_k) = \sum_{k=1}^{\infty} a_k + \sum_{k=1}^{\infty} b_k. \tag{63}$$

If c is any constant then

$$ca_1 + ca_2 + ca_3 + \cdots = c(a_1 + a_2 + a_3 + \cdots) \quad \textit{i.e.} \quad \sum_{k=1}^{\infty} ca_k = c\sum_{k=1}^{\infty} a_k.$$

Another way to write equation (63) is

$$\begin{aligned}
&(a_1 + b_1) + (a_2 + b_2) + (a_3 + b_3) + \cdots \\
&\qquad\qquad = (a_1 + a_2 + a_3 + \cdots) + (b_1 + b_2 + b_3 + \cdots).
\end{aligned}$$

Rearranging terms in a series. Note that to get the left hand side from the right hand side we have to switch around infinitely many terms! This may not seem very surprising, but it turns out that if one digs deeper into the subject of series, examples of series show up where switching infinitely many terms around actually changes the sum.

A simplest example of this is the sum $1 - 1 + 1 - 1 \ldots$ from the introduction. Here we could first try to add all the positive terms $(1 + 1 + 1 + \cdots)$ and then the negative terms $(-1 - 1 - 1 \cdots)$, and finally combine them:

$$1 - 1 + 1 - 1 + 1 - 1 = (1 + 1 + 1 + \cdots) + (-1 - 1 - 1 \cdots) = \infty - \infty?$$

This clearly does not make a lot of sense. But, since the series in this example does not converge, that was perhaps to be expected.

The following example however presents a convergent series whose sum changes if one rearranges the terms. If we take Leibniz' formula for $\ln 2$ (obtained from the Taylor expansion for $\ln(1 + x)$; see §7 below),

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \cdots = \ln 2 \tag{64}$$

and rearrange the terms so that we add two positive terms and then one negative term at a time we get

$$1+\frac{1}{3}-\frac{1}{2}+\frac{1}{5}+\frac{1}{7}-\frac{1}{4}+\frac{1}{9}+\frac{1}{11}-\frac{1}{6}+\cdots=\frac{3}{2}\ln 2 \tag{65}$$

This fact is not obvious; the explanation would take us beyond the scope of this course, although it does not require any ideas that a student in math 222 would not be familiar with.

5. Convergence of Taylor Series

We now return to the study of Taylor series and their convergence.

Let $y = f(x)$ be some function defined on an interval $a < x < b$ containing 0. For any number x the function defines a series

$$f(0)+f'(0)x+\frac{f''(0)}{2!}x^2+\frac{f^{(3)}(0)}{3!}x^3+\cdots \tag{66}$$

which is called the Taylor series of f. This raises two questions:

- Does the Taylor series converge?
- If the Taylor series converges, then what is its sum?

Since each different choice of x leads to a different series, the answer to these questions depends on x.

There is no easy and general criterion that we could apply to a given function $f(x)$ to find out if its Taylor series converges for any particular x (except $x = 0$ – what does the Taylor series look like when we set $x = 0$?). On the other hand, it turns out that for many functions the Taylor series does converge to $f(x)$ for all x in some interval $-\rho < x < \rho$. In this section we will check this for two examples: the "geometric series" and the exponential function.

But first, before we do the examples, a word about how we will prove that Taylor series converges: instead of taking the limit of the $T_nf(x)$ as $n \to \infty$, we are usually better off looking at the remainder term. Since $T_nf(x) = f(x) - R_nf(x)$ we have

$$\lim_{n\to\infty} T_nf(x) = f(x) \iff \lim_{n\to\infty} R_nf(x) = 0$$

So, to check that the Taylor series of $f(x)$ converges to $f(x)$ we must show that the remainder term $R_nf(x)$ goes to zero as $n \to \infty$.

5.1. The Geometric series converges for $-1 < x < 1$.

If $f(x) = 1/(1 - x)$ then by the formula for the Geometric Sum we have

$$\begin{aligned} f(x) &= \frac{1}{1-x} \\ &= \frac{1-x^{n+1}+x^{n+1}}{1-x} \\ &= 1+x+x^2+\cdots+x^n+\frac{x^{n+1}}{1-x} \\ &= T_nf(x)+\frac{x^{n+1}}{1-x}. \end{aligned}$$

We are not dividing by zero since $|x| < 1$ so that $1 - x \neq 0$. The remainder term is

$$R_nf(x) = \frac{x^{n+1}}{1-x}.$$

Since $|x| < 1$ we have

$$\lim_{n\to\infty} |R_n f(x)| = \lim_{n\to\infty} \frac{|x|^{n+1}}{|1-x|} = \frac{\lim_{n\to\infty} |x|^{n+1}}{|1-x|} = \frac{0}{|1-x|} = 0.$$

Thus we have shown that the series converges for all $-1 < x < 1$, i.e.

$$\frac{1}{1-x} = \lim_{n\to\infty} \left\{1 + x + x^2 + \cdots + x^n\right\} = 1 + x + x^2 + x^3 + \cdots = \sum_{k=0}^{\infty} x^k.$$

5.2. Convergence of the exponential Taylor series. Let $f(x) = e^x$. It turns out the Taylor series of e^x converges to e^x for every value of x. Here's why: we had found that

$$T_n e^x = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!},$$

and by Lagrange's formula the remainder is given by

$$R_n e^x = e^{\xi} \frac{x^{n+1}}{(n+1)!},$$

where ξ is some number between 0 and x.

If $x > 0$ then $0 < \xi < x$ so that $e^{\xi} \le e^x$; if $x < 0$ then $x < \xi < 0$ implies that $e^{\xi} < e^0 = 1$. Either way one has $e^{\xi} \le e^{|x|}$, and thus

$$|R_n e^x| \le e^{|x|} \frac{|x|^{n+1}}{(n+1)!}.$$

We have shown before that $\lim_{n\to\infty} x^{n+1}/(n+1)! = 0$, so the Sandwich theorem again implies that $\lim_{n\to\infty} |R_n e^x| = 0$.

Conclusion:

$$e^x = \lim_{n\to\infty} \left\{1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!}\right\} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots$$

Do Taylor series always converge? And if the series of some function $y = f(x)$ converges, must it then converge to $f(x)$? Although the Taylor series of almost any function we run into will converge to the function itself, the following example shows that it doesn't have to be so.

5.3. The day that all Chemistry stood still. The rate at which a chemical reaction "A→B" proceeds depends among other things on the temperature at which the reaction is taking place. This dependence is described by the ***Arrhenius law*** which states that the rate at which a reaction takes place is proportional to

$$f(T) = e^{-\frac{\Delta E}{kT}}$$

where ΔE is the amount of energy involved in each reaction, k is Boltzmann's constant, and T is the temperature in degrees Kelvin. If we ignore the constants ΔE and k (i.e. if we set $\Delta E/k$ equal to one by choosing the right units) then the reaction rate is proportional to

$$f(T) = e^{-1/T}.$$

If we have to deal with reactions at low temperatures we might be inclined to replace this function with its Taylor series at $T = 0$, or at least the first non-zero term in this series.

If we were to do this we would be in for a surprise. To see what happens, let's look at the following function,

$$f(x) = \begin{cases} e^{-1/x} & x > 0 \\ 0 & x \le 0 \end{cases}$$

This function goes to zero ***very*** quickly as $x \to 0$. In fact one has

$$\lim_{x \searrow 0} \frac{f(x)}{x^n} = \lim_{x \searrow 0} \frac{e^{-1/x}}{x^n} = \lim_{t \to \infty} t^n e^{-t} = 0. \qquad \text{(set } t = 1/x\text{)}$$

This implies

$$f(x) = o(x^n) \quad (x \to 0)$$

for any $n = 1, 2, 3 \ldots$. As $x \to 0$, this function vanishes faster than any power of x.

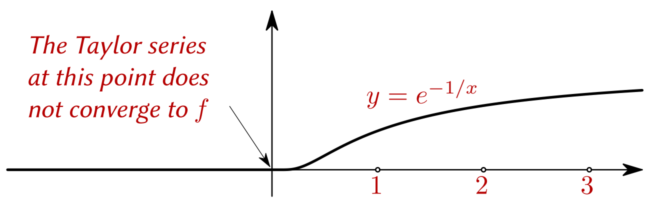


Figure 1. An innocent looking function with an unexpected Taylor series. See example 5.3 which shows that even when a Taylor series of some function f converges we can't be sure that it converges to f – it could converge to a different function.

If we try to compute the Taylor series of f we need its derivatives at $x = 0$ of all orders. These can be computed (not easily), and the result turns out to be that ***all derivatives of f vanish at $x = 0$,***

$$f(0) = f'(0) = f''(0) = f^{(3)}(0) = \cdots = 0.$$

The Taylor series of f is therefore

$$T_\infty f(x) = 0 + 0 \cdot x + 0 \cdot \frac{x^2}{2!} + 0 \cdot \frac{x^3}{3!} + \cdots = 0.$$

Clearly this series converges (all terms are zero, after all), but instead of converging to the function $f(x)$ we started with, it converges to the function $g(x) = 0$.

What does this mean for the chemical reaction rates and Arrhenius' law? We wanted to "simplify" the Arrhenius law by computing the Taylor series of $f(T)$ at $T = 0$, but we have just seen that all terms in this series are zero. Therefore replacing the Arrhenius reaction rate by its Taylor series at $T = 0$ has the effect of setting all reaction rates equal to zero.

6. Problems on Convergence of Taylor Series

1. Prove that the Taylor series for $f(x) = \cos x$ converges to $f(x)$ for all real numbers x (by showing that the remainder term goes to zero as $n \to \infty$).

2. Prove that the Taylor series for $g(x) = \sin(2x)$ converges to $g(x)$ for all real numbers x .

3. Prove that the Taylor series for $h(x) = \cosh(x)$ converges to $h(x)$ for all real numbers x .

4. Prove that the Taylor series for $k(x) = e^{2x+3}$ converges to $k(x)$ for all real numbers x .

5. Prove that the Taylor series for $\ell(x) = \cos\bigl(x - \frac{\pi}{7}\bigr)$ converges to $\ell(x)$ for all real numbers x.

6. **[Group Problem]** If the Taylor series of a function $y = f(x)$ converges for all x, does it have to converge to $f(x)$, or could it converge to some other function?

7. For which real numbers x does the Taylor series of $f(x) = \dfrac{1}{1-x}$ converge to $f(x)$?

8. For which real numbers x does the Taylor series of $f(x) = \dfrac{1}{1-x^2}$ converge to $f(x)$? (hint: a substitution may help.)

9. For which real numbers x does the Taylor series of $f(x) = \dfrac{1}{1+x^2}$ converge to $f(x)$?

10. For which real numbers x does the Taylor series of $f(x) = \dfrac{1}{3+2x}$ converge to $f(x)$?

11. For which real numbers x does the Taylor series of $f(x) = \frac{1}{2-5x}$ converge to $f(x)$?

12. **[Group Problem]** For which real numbers x does the Taylor series of $f(x) = \dfrac{1}{2-x-x^2}$ converge to $f(x)$? (hint: use PFD and the Geometric Series to find the remainder term.)

13. Show that the Taylor series for $f(x) = \ln(1+x)$ converges when $-1 < x < 1$ by integrating the Geometric Series

$$\frac{1}{1+t} = 1 - t + t^2 - t^3 + \cdots + (-1)^n t^n + (-1)^{n+1}\frac{t^{n+1}}{1+t}$$

from $t = 0$ to $t = x$. (See §7.)

14. Show that the Taylor series for $f(x) = e^{-x^2}$ converges for all real numbers x. (Set $t = -x^2$ in the Taylor series with remainder for e^t.)

15. Show that the Taylor series for $f(x) = \sin(x^4)$ converges for all real numbers x. (Set $t = x^4$ in the Taylor series with remainder for $\sin t$.)

16. Show that the Taylor series for $f(x) = 1/(1+x^3)$ converges whenever $-1 < x < 1$ (Use the GEOMETRIC SERIES.)

17. For which x does the Taylor series of $f(x) = 2/(1+4x^2)$ converge? (Again, use the GEOMETRIC SERIES.)

7. Leibniz' formulas for $\ln 2$ and $\pi/4$

Leibniz showed that

$$\frac{1}{1} - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \cdots = \ln 2$$

and

$$\frac{1}{1} - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \cdots = \frac{\pi}{4}$$

Both formulas arise by setting $x = 1$ in the Taylor series for

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} + \frac{x^4}{4} - \cdots$$
$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} + \frac{x^7}{7} - \cdots$$

This is only justified if we show that the series actually converge, which we'll do here, at least for the first of these two formulas. The proof of the second is similar. The following is not Leibniz' original proof.

We begin with the geometric sum

$$1 - x + x^2 - x^3 + \cdots + (-1)^n x^n = \frac{1}{1+x} + \frac{(-1)^{n+1}x^{n+1}}{1+x}$$

Then we integrate both sides from $x = 0$ to $x = 1$ and get

$$\begin{aligned}\frac{1}{1} - \frac{1}{2} + \frac{1}{3} - \cdots + (-1)^n \frac{1}{n+1} &= \int_0^1 \frac{\mathrm{d}x}{1+x} + (-1)^{n+1} \int_0^1 \frac{x^{n+1}\mathrm{d}x}{1+x} \\ &= \ln 2 + (-1)^{n+1} \int_0^1 \frac{x^{n+1}\mathrm{d}x}{1+x}\end{aligned}$$

(Use $\int_0^1 x^k \mathrm{d}x = \frac{1}{k+1}$.) Instead of computing the last integral we estimate it by saying

$$0 \le \frac{x^{n+1}}{1+x} \le x^{n+1} \implies 0 \le \int_0^1 \frac{x^{n+1}\mathrm{d}x}{1+x} \le \int_0^1 x^{n+1}\mathrm{d}x = \frac{1}{n+2}$$

Hence

$$\lim_{n\to\infty} (-1)^{n+1} \int_0^1 \frac{x^{n+1}\mathrm{d}x}{1+x} = 0,$$

and we get

$$\lim_{n\to\infty} \frac{1}{1} - \frac{1}{2} + \frac{1}{3} - \cdots + (-1)^n \frac{1}{n+1} = \ln 2 + \lim_{n\to\infty} (-1)^{n+1} \int_0^1 \frac{x^{n+1}\mathrm{d}x}{1+x} = \ln 2.$$

8. Problems

1. [Group Problem] The error function from statistics is defined by

$$\mathrm{erf}(x) = \frac{1}{\sqrt{\pi}} \int_0^x e^{-t^2/2}\,\mathrm{d}t$$

(a) Find the Taylor series of the error function from the Taylor series of $f(r) = e^r$ (set $r = -t^2/2$ and integrate).

(b) Estimate the error term and show that the Taylor series of the error function converges for all real x.

2. [Group Problem] Prove Leibniz' formula for $\frac{\pi}{4}$ by mimicking the proof in section 7. Specifically, find a formula for the remainder in :

$$\frac{1}{1+t^2} = 1 - t^2 + \cdots + (-1)^n t^{2n} + R_{2n}(t)$$

and integrate this from $t = 0$ to $t = 1$.

CHAPTER 6

Vectors

1. Introduction to vectors

1.1. Definition. *A vector is a column of two, three, or more numbers, written as*

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \quad \text{or} \quad \vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \quad \text{or} \quad \vec{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}.$$

*The **length of a vector** $\vec{a} = \left(\begin{smallmatrix} a_1 \\ a_2 \\ a_3 \end{smallmatrix}\right)$ is defined by*

$$\|\vec{a}\| = \left\| \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \right\| = \sqrt{a_1^2 + a_2^2 + a_3^2}.$$

We will always deal with either the two or three dimensional cases, in other words, the cases $n = 2$ or $n = 3$, respectively. For these cases there is a geometric description of vectors which is very useful. In fact, the two and three dimensional theories have their origins in physics and geometry. In higher dimensions the geometric description fails, simply because we cannot visualize a four dimensional space, let alone a higher dimensional space. Instead of a geometric description of vectors there is an abstract theory called ***Linear Algebra*** which deals with "vector spaces" of any dimension (even infinite!). This theory of vectors in higher dimensional spaces is very useful in science, engineering and economics. You can learn about it in courses like MATH 320 or 340/341.

1.2. Basic arithmetic of vectors. You can add and subtract vectors, and you can multiply them with arbitrary real numbers. this section tells you how.

The ***sum of two vectors*** is defined by

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \end{pmatrix}, \tag{67}$$

and

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ a_3 + b_3 \end{pmatrix}.$$

The ***zero vector*** is defined by

$$\vec{0} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{or} \quad \vec{0} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

It has the property that

$$\vec{a} + \vec{0} = \vec{0} + \vec{a} = \vec{a}$$

no matter what the vector $\vec{a}$ is.

You can multiply a vector $\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ with a real number t according to the rule

$$t\vec{a} = \begin{pmatrix} ta_1 \\ ta_2 \\ ta_3 \end{pmatrix}.$$

In particular, "minus a vector" is defined by

$$-\vec{a} = (-1)\vec{a} = \begin{pmatrix} -a_1 \\ -a_2 \\ -a_3 \end{pmatrix}.$$

The difference of two vectors is defined by

$$\vec{a} - \vec{b} = \vec{a} + (-\vec{b}).$$

So, to subtract two vectors you subtract their components,

$$\vec{a} - \vec{b} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} - \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_1 - b_1 \\ a_2 - b_2 \\ a_3 - b_3 \end{pmatrix}$$

1.3. Some GOOD examples.

$$\begin{pmatrix} 2 \\ 3 \end{pmatrix} + \begin{pmatrix} -3 \\ \pi \end{pmatrix} = \begin{pmatrix} -1 \\ 3+\pi \end{pmatrix} \qquad 2\begin{pmatrix} 1 \\ 0 \end{pmatrix} + 3\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

$$\left(\begin{smallmatrix} 1 \\ 0 \\ 3 \end{smallmatrix}\right) - \left(\begin{smallmatrix} -1 \\ 12 \\ \sqrt{2} \end{smallmatrix}\right) = \left(\begin{smallmatrix} 2 \\ -12 \\ 3-\sqrt{2} \end{smallmatrix}\right) \qquad a\left(\begin{smallmatrix} 1 \\ 0 \\ 0 \end{smallmatrix}\right) + b\left(\begin{smallmatrix} 0 \\ 1 \\ 0 \end{smallmatrix}\right) + c\left(\begin{smallmatrix} 0 \\ 0 \\ 1 \end{smallmatrix}\right) = \left(\begin{smallmatrix} a \\ b \\ c \end{smallmatrix}\right)$$

$$0 \cdot \begin{pmatrix} 12\sqrt{39} \\ \pi^2 - \ln 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \vec{0} \qquad \begin{pmatrix} t+t^2 \\ 1-t^2 \end{pmatrix} = (1+t)\begin{pmatrix} t \\ 1-t \end{pmatrix}$$

1.4. Two very, very BAD examples. Vectors must have the same size to be added, therefore

$$\begin{pmatrix} 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} = \textbf{undefined!!!}$$

Vectors and numbers are different things, so an equation like

$$\vec{a} = 3 \quad \text{is nonsense!}$$

This equation says that some vector ($\vec{a}$) is equal to some number (in this case: 3). ***Vectors and numbers are never equal!***

1.5. Algebraic properties of vector addition and multiplication. Addition of vectors and multiplication of numbers and vectors were defined in such a way that the following always hold for any vectors $\vec{a}$, $\vec{b}$, $\vec{c}$ (of the same size) and any real numbers s, t

(68a) $$\vec{a} + \vec{b} = \vec{b} + \vec{a}$$ [vector addition is commutative]

(68b) $$\vec{a} + (\vec{b} + \vec{c}) = (\vec{a} + \vec{b}) + \vec{c}$$ [vector addition is associative]

(68c) $$t(\vec{a} + \vec{b}) = t\vec{a} + t\vec{b}$$ [first distributive property]

(68d) $$(s+t)\vec{a} = s\vec{a} + t\vec{a}$$ [second distributive property]

Prove (68a). Let $\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ and $\vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ be two vectors, and consider both possible ways of adding them:

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ a_3 + b_3 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} + \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} b_1 + a_1 \\ b_2 + a_2 \\ b_3 + a_3 \end{pmatrix}$$

We know (or we have assumed long ago) that addition of real numbers is commutative, so that $a_1 + b_1 = b_1 + a_1$, etc. Therefore

$$\vec{a} + \vec{b} = \begin{pmatrix} a_1+b_1 \\ a_2+b_2 \\ a_3+b_3 \end{pmatrix} = \begin{pmatrix} b_1+a_1 \\ b_2+a_2 \\ b_3+a_3 \end{pmatrix} = \vec{b} + \vec{a}.$$

This proves (68a).

The properties (68a), ..., (68d), are used to facilitate the manipulation of vectors. The following are examples of how the arithmetic of vectors works.

1.6. Example. If $\vec{v}$ and $\vec{w}$ are two vectors, we define

$$\vec{a} = 2\vec{v} + 3\vec{w}, \quad \vec{b} = -\vec{v} + \vec{w}.$$

1.6.1. *Problem.* Compute $\vec{a} + \vec{b}$ and $2\vec{a} - 3\vec{b}$ in terms of $\vec{v}$ and $\vec{w}$.
Solution:

$$\vec{a} + \vec{b} = (2\vec{v} + 3\vec{w}) + (-\vec{v} + \vec{w}) = (2-1)\vec{v} + (3+1)\vec{w} = \vec{v} + 4\vec{w}$$
$$2\vec{a} - 3\vec{b} = 2(2\vec{v} + 3\vec{w}) - 3(-\vec{v} + \vec{w}) = 4\vec{w} + 6\vec{w} + 3\vec{v} - 3\vec{w} = 7\vec{v} + 3\vec{w}.$$

1.6.2. *Problem.* Find s, t so that $s\vec{a} + t\vec{b} = \vec{v}$.
Solution: Simplifying $s\vec{a} + t\vec{b}$ you find

$$s\vec{a} + t\vec{b} = s(2\vec{v} + 3\vec{w}) + t(-\vec{v} + \vec{w}) = (2s - t)\vec{v} + (3s + t)\vec{w}.$$

One way to ensure that $s\vec{a} + t\vec{b} = \vec{v}$ holds is therefore to choose s and t to be the solutions of

$$2s - t = 1$$
$$3s + t = 0$$

The second equation says $t = -3s$. The first equation then leads to $2s + 3s = 1$, i.e. $s = \frac{1}{5}$. Since $t = -3s$ we get $t = -\frac{3}{5}$. The solution we have found is therefore

$$\tfrac{1}{5}\vec{a} - \tfrac{3}{5}\vec{b} = \vec{v}.$$

2. Geometric description of vectors

Vectors originally appeared in physics, where they represented forces: a force acting on some object has a ***magnitude*** and a ***direction***. Thus a force can be thought of as an arrow in the direction of the force, where the length of the arrow indicates how strong the force is (how hard it pushes or pulls).

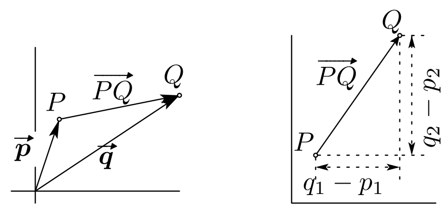
So we will think of vectors as ***arrows***: if you specify two points P and Q, then the arrow pointing from P to Q is a vector and we denote this vector by $\overrightarrow{PQ}$.

The precise mathematical definition is as follows:

2.1. Definition. *For any pair of points P and Q whose coordinates are (p_1, p_2, p_3) and (q_1, q_2, q_3) one defines a vector $\overrightarrow{PQ}$ by*

$$\overrightarrow{PQ} = \begin{pmatrix} q_1 - p_1 \\ q_2 - p_2 \\ q_3 - p_3 \end{pmatrix}.$$

*If the initial point of an arrow is the origin O, and the final point is any point Q, then the vector $\overrightarrow{OQ}$ is called the **position vector** of the point Q.*



If $\vec{\boldsymbol{p}}$ and $\vec{\boldsymbol{q}}$ are the position vectors of P and Q, then one can write $\overrightarrow{PQ}$ as

$$\overrightarrow{PQ} = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} - \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} = \vec{\boldsymbol{q}} - \vec{\boldsymbol{p}}.$$

For plane vectors we define $\overrightarrow{PQ}$ similarly, namely, $\overrightarrow{PQ} = \left(\begin{smallmatrix} q_1 - p_1 \\ q_2 - p_2 \end{smallmatrix}\right)$. The old formula for the distance between two points P and Q in the plane

$$\text{distance from } P \text{ to } Q = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2}$$

says that the length of the vector $\overrightarrow{PQ}$ is just the distance between the points P and Q, i.e.

$$\text{distance from } P \text{ to } Q = \left\| \overrightarrow{PQ} \right\|.$$

This formula is also valid if P and Q are points in space, in which case it is given by

$$\text{distance from } P \text{ to } Q = \sqrt{(q_1 - p_1)^2 + (q_2 - p_2)^2 + (q_3 - p_3)^2}.$$

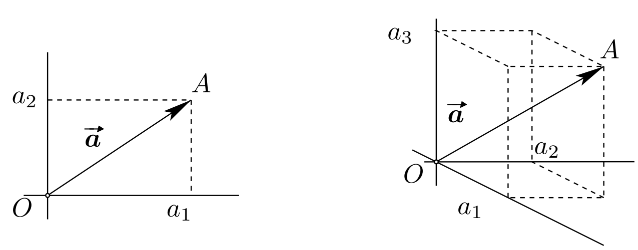


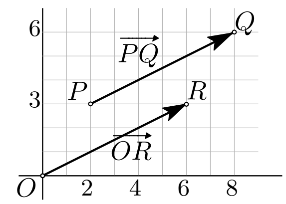
Figure 1. Position vectors in the plane and in space

2.2. Example. The point P has coordinates $(2,3)$; the point Q has coordinates $(8,6)$. The vector $\overrightarrow{PQ}$ is therefore

$$\overrightarrow{PQ} = \begin{pmatrix} 8-2 \\ 6-3 \end{pmatrix} = \begin{pmatrix} 6 \\ 3 \end{pmatrix}.$$

This vector is the position vector of the point R whose coordinates are $(6,3)$. Thus

$$\overrightarrow{PQ} = \overrightarrow{OR} = \begin{pmatrix} 6 \\ 3 \end{pmatrix}.$$



The distance from P to Q is the length of the vector $\overrightarrow{PQ}$, i.e.

$$\text{distance } P \text{ to } Q = \left\| \begin{pmatrix} 6 \\ 3 \end{pmatrix} \right\| = \sqrt{6^2+3^2} = 3\sqrt{5}.$$

2.3. Example. Find the distance between the points A and B whose position vectors are $\vec{a} = \left(\begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix}\right)$ and $\vec{b} = \left(\begin{smallmatrix} 0 \\ 1 \\ 1 \end{smallmatrix}\right)$ respectively.

Solution: One has

$$\text{distance } A \text{ to } B = \|\overrightarrow{AB}\| = \|\vec{b} - \vec{a}\| = \left\| \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\| = \sqrt{(-1)^2+0^2+1^2} = \sqrt{2}$$

2.4. Geometric interpretation of vector addition and multiplication. Suppose you have two vectors $\vec{a}$ and $\vec{b}$. Consider them as position vectors, i.e. represent them by vectors that have the origin as initial point:

$$\vec{a} = \overrightarrow{OA}, \quad \vec{b} = \overrightarrow{OB}.$$

Then the origin and the three endpoints of the vectors $\vec{a}$, $\vec{b}$ and $\vec{a}+\vec{b}$ form a parallelogram. See figure 2.

To multiply a vector $\vec{a}$ with a real number t you multiply its length with $|t|$; if $t<0$ you reverse the direction of $\vec{a}$.

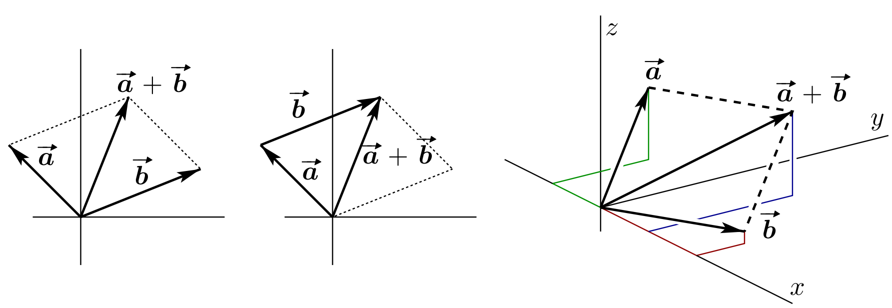


Figure 2. Two ways of adding plane vectors, and an addition of space vectors

2.5. Example. In example 1.6 we assumed two vectors $\vec{v}$ and $\vec{w}$ were given, and then defined $\vec{a} = 2\vec{v}+3\vec{w}$ and $\vec{b} = -\vec{v}+\vec{w}$. In figure 4 the vectors $\vec{a}$ and $\vec{b}$ are constructed geometrically from some arbitrarily chosen $\vec{v}$ and $\vec{w}$. We also found algebraically in example 1.6 that $\vec{a}+\vec{b} = \vec{v}+4\vec{w}$. The third drawing in figure 4 illustrates this.

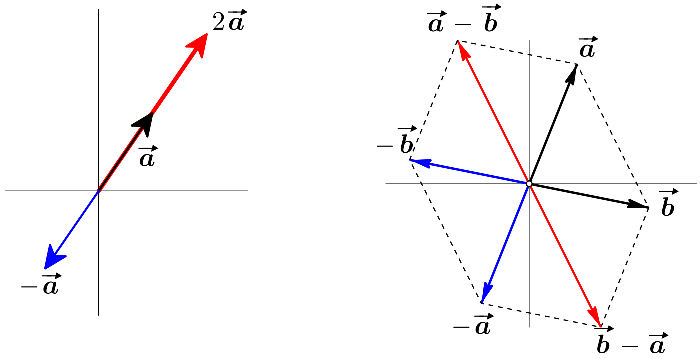


Figure 3. Multiples of a vector, and the difference of two vectors.

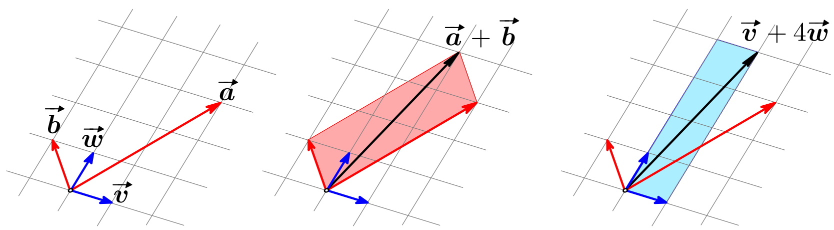
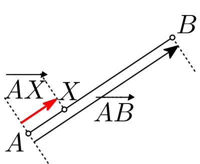


Figure 4. Picture proof that $\vec{a}+\vec{b}=\vec{v}+4\vec{w}$ in example 2.5.

3. Parametric equations for lines and planes



Given two ***distinct*** points A and B we consider the line segment AB. If X is any given point on AB then we will now find a formula for the position vector of X.

Define t to be the ratio between the lengths of the line segments AX and AB,

$$t = \frac{\text{length } AX}{\text{length } AB}.$$

Then the vectors $\overrightarrow{AX}$ and $\overrightarrow{AB}$ are related by $\overrightarrow{AX} = t\overrightarrow{AB}$. Since AX is shorter than AB we have $0 < t < 1$.

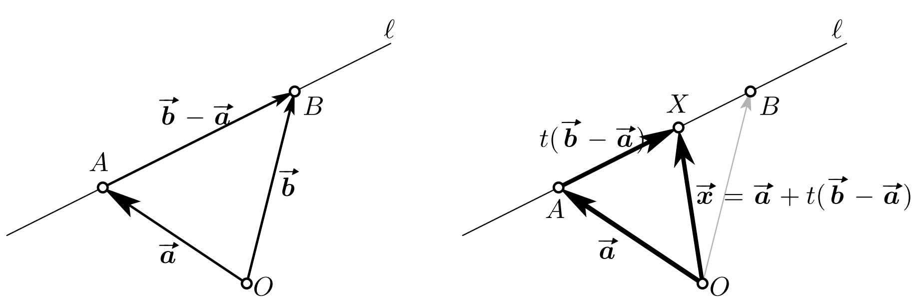


Figure 5. Constructing points on the line through A and B

The position vector of the point X on the line segment AB is

$$\overrightarrow{OX} = \overrightarrow{OA} + \overrightarrow{AX} = \overrightarrow{OA} + t\overrightarrow{AB}.$$

If we write $\vec{a}, \vec{b}, \vec{x}$ for the position vectors of A, B, X, then we get $\overrightarrow{AX} = \vec{x} - \vec{a}$ and $\overrightarrow{AB} = \vec{b} - \vec{a}$, so that

$$\vec{x} = \vec{a} + t(\vec{b} - \vec{a}). \tag{69}$$

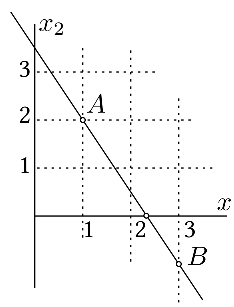
which is sometimes also written as

$$\vec{x} = (1-t)\vec{a} + t\vec{b}.$$

This equation is called the ***parametric equation for the line through*** A ***and*** B.

In our derivation the parameter t satisfied $0 \le t \le 1$, but there is nothing that keeps us from substituting negative values of t, or numbers $t > 1$ in (69). The resulting vectors $\vec{x}$ are position vectors of points X which lie on the line ℓ through A and B.

3.1. Example. *Find the parametric equation for the line ℓ through the points $A(1,2)$ and $B(3,-1)$, and determine where ℓ intersects the x_1 axis.*



Solution: The position vectors of A, B are $\vec{a} = \binom{1}{2}$ and $\vec{b} = \binom{3}{-1}$, so the position vector of any point on ℓ is given by

$$\vec{x} = \vec{a} + t(\vec{b} - \vec{a}) = \begin{pmatrix}1\\2\end{pmatrix} + t\begin{pmatrix}3-1\\-1-2\end{pmatrix} = \begin{pmatrix}1\\2\end{pmatrix} + t\begin{pmatrix}2\\-3\end{pmatrix} = \begin{pmatrix}1+2t\\2-3t\end{pmatrix}$$

where t is an arbitrary real number.

This vector points to the point $X = (1+2t, 2-3t)$. By definition, a point lies on the x_1-axis if its x_2 component vanishes. Thus if the point

$$X = (1+2t, 2-3t)$$

lies on the x_1-axis, then $2 - 3t = 0$, i.e. $t = \frac{2}{3}$. When $t = \frac{2}{3}$ the x_1-coordinate of X is $1 + 2t = \frac{5}{3}$, so the intersection point ℓ and the x_1-axis is $X = (1 + 2\cdot\frac{2}{3}, 0) = (\frac{5}{3}, 0)$.

3.2. Midpoint of a line segment. If M is the midpoint of the line segment AB, then the vectors $\overrightarrow{AM}$ and $\overrightarrow{MB}$ are both parallel and have the same direction and length (namely, half the length of the line segment AB). Hence they are equal: $\overrightarrow{AM} = \overrightarrow{MB}$. If $\vec{a}$, $\vec{m}$, and $\vec{b}$ are the position vectors of A , M and B, then this means

$$\vec{m} - \vec{a} = \overrightarrow{AM} = \overrightarrow{MB} = \vec{b} - \vec{m}.$$

Add $\vec{m}$ and $\vec{a}$ to both sides, and divide by 2 to get

$$\vec{m} = \tfrac{1}{2}\vec{a} + \tfrac{1}{2}\vec{b} = \frac{\vec{a} + \vec{b}}{2}.$$

4. Vector Bases

4.1. The Standard Basis Vectors. The notation for vectors which we have been using so far is not the most traditional. In the late 19th century Gibbs and Heavyside adapted Hamilton's theory of Quaternions to deal with vectors. Their notation is still popular in texts on electromagnetism and fluid mechanics.

Define the following three vectors:

$$\vec{\imath} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \vec{\jmath} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{k} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Then *every vector can be written as a linear combination of* $\vec{\imath}$, $\vec{\jmath}$ *and* $\vec{k}$*,* namely as follows:

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = a_1\vec{\imath} + a_2\vec{\jmath} + a_3\vec{k}.$$

Moreover, *there is only one way to write a given vector as a linear combination of* $\{\vec{\imath}, \vec{\jmath}, \vec{k}\}$. This means that

$$a_1\vec{\imath} + a_2\vec{\jmath} + a_3\vec{k} = b_1\vec{\imath} + b_2\vec{\jmath} + b_3\vec{k} \iff \begin{cases} a_1 = b_1 \\ a_2 = b_2 \\ a_3 = b_3 \end{cases}$$

For plane vectors one defines

$$\vec{\imath} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \vec{\jmath} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

and just as for three dimensional vectors one can write every (plane) vector $\vec{a}$ as a linear combination of $\vec{\imath}$ and $\vec{\jmath}$,

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = a_1\vec{\imath} + a_2\vec{\jmath}.$$

Just as for space vectors, there is only one way to write a given vector as a linear combination of $\vec{\imath}$ and $\vec{\jmath}$.

4.2. A Basis of Vectors (in general)*. The vectors $\vec{\imath}, \vec{\jmath}, \vec{k}$ are called the ***standard basis vectors***. They are an example of what is called a "basis". Here is the definition in the case of space vectors:

4.3. Definition. *A triple of space vectors* $\{\vec{u}, \vec{v}, \vec{w}\}$ *is a* ***basis*** *if every space vector* $\vec{a}$ *can be written as a linear combination of* $\{\vec{u}, \vec{v}, \vec{w}\}$*, i.e.*

$$\vec{a} = a\vec{u} + b\vec{v} + c\vec{w},$$

and if there is only one way to do so for any given vector $\vec{a}$ *(i.e. the vector* $\vec{a}$ *determines the coefficients* a, b, c*).* For plane vectors the definition of a basis is almost the same, except that a basis consists of two vectors rather than three:

4.4. Definition. *A pair of plane vectors* $\{\vec{u}, \vec{v}\}$ *is a* ***basis*** *if every plane vector* $\vec{a}$ *can be written as a linear combination of* $\{\vec{u}, \vec{v}\}$*, i.e.* $\vec{a} = a\vec{u} + b\vec{v}$*, and if there is only one way to do so for any given vector* $\vec{a}$ *(i.e. the vector* $\vec{a}$ *determines the coefficients* a, b*).*

5. Dot Product

5.1. Definition. *The "inner product" or "dot product" of two vectors is given by*

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = a_1b_1 + a_2b_2 + a_3b_3.$$

Note that the dot-product of two vectors is a number!

The dot product of two plane vectors is (predictably) defined by

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = a_1b_1 + a_2b_2.$$

An important property of the dot product is its relation with the length of a vector:

$$\|\vec{\boldsymbol{a}}\|^2 = \vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{a}}. \tag{70}$$

5.2. Algebraic properties of the dot product. The dot product satisfies the following rules,

$$\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{b}} = \vec{\boldsymbol{b}} \cdot \vec{\boldsymbol{a}} \tag{71}$$

$$\vec{\boldsymbol{a}} \cdot (\vec{\boldsymbol{b}} + \vec{\boldsymbol{c}}) = \vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{b}} + \vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{c}} \tag{72}$$

$$(\vec{\boldsymbol{b}} + \vec{\boldsymbol{c}}) \cdot \vec{\boldsymbol{a}} = \vec{\boldsymbol{b}} \cdot \vec{\boldsymbol{a}} + \vec{\boldsymbol{c}} \cdot \vec{\boldsymbol{a}} \tag{73}$$

$$t(\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{b}}) = (t\vec{\boldsymbol{a}}) \cdot \vec{\boldsymbol{b}} \tag{74}$$

which hold for all vectors $\vec{\boldsymbol{a}}$, $\vec{\boldsymbol{b}}$, $\vec{\boldsymbol{c}}$ and any real number t.

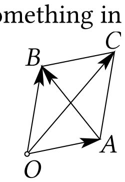
As usual, these properties allow us to develop the arithmetic of the dot product, as in the next example.

5.3. Example. Simplify $\|\vec{\boldsymbol{a}} + \vec{\boldsymbol{b}}\|^2$.

One has

$$\begin{aligned} \|\vec{\boldsymbol{a}} + \vec{\boldsymbol{b}}\|^2 &= (\vec{\boldsymbol{a}} + \vec{\boldsymbol{b}}) \cdot (\vec{\boldsymbol{a}} + \vec{\boldsymbol{b}}) \\ &= \vec{\boldsymbol{a}} \cdot (\vec{\boldsymbol{a}} + \vec{\boldsymbol{b}}) + \vec{\boldsymbol{b}} \cdot (\vec{\boldsymbol{a}} + \vec{\boldsymbol{b}}) \\ &= \vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{a}} + \underbrace{\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{b}} + \vec{\boldsymbol{b}} \cdot \vec{\boldsymbol{a}}}_{=2\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{b}} \text{ by (71)}} + \vec{\boldsymbol{b}} \cdot \vec{\boldsymbol{b}} \\ &= \|\vec{\boldsymbol{a}}\|^2 + 2\vec{\boldsymbol{a}} \cdot \vec{\boldsymbol{b}} + \|\vec{\boldsymbol{b}}\|^2 \end{aligned}$$

5.4. The diagonals of a parallelogram. Here is an example of how you can use the algebra of the dot product to prove something in geometry.



Suppose you have a parallelogram one of whose vertices is the origin. Label the vertices, starting at the origin and going around counterclockwise, O, A, C and B. Let $\vec{\boldsymbol{a}} = \overrightarrow{OA}$, $\vec{\boldsymbol{b}} = \overrightarrow{OB}$, $\vec{\boldsymbol{c}} = \overrightarrow{OC}$. One has

$$\overrightarrow{OC} = \vec{\boldsymbol{c}} = \vec{\boldsymbol{a}} + \vec{\boldsymbol{b}}, \quad \text{and} \quad \overrightarrow{AB} = \vec{\boldsymbol{b}} - \vec{\boldsymbol{a}}.$$

These vectors correspond to the diagonals OC and AB

5.5. Theorem. *In a parallelogram $OACB$ the sum of the squares of the lengths of the two diagonals equals the sum of the squares of the lengths of all four sides.*

PROOF. The squared lengths of the diagonals are

$$\|\overrightarrow{OC}\|^2 = \|\vec{\boldsymbol a}+\vec{\boldsymbol b}\|^2 = \|\vec{\boldsymbol a}\|^2 + 2\vec{\boldsymbol a}\bullet\vec{\boldsymbol b} + \|\vec{\boldsymbol b}\|^2$$
$$\|\overrightarrow{AB}\|^2 = \|\vec{\boldsymbol a}-\vec{\boldsymbol b}\|^2 = \|\vec{\boldsymbol a}\|^2 - 2\vec{\boldsymbol a}\bullet\vec{\boldsymbol b} + \|\vec{\boldsymbol b}\|^2$$

Adding both these equations you get

$$\|\overrightarrow{OC}\|^2 + \|\overrightarrow{AB}\|^2 = 2\left(\|\vec{\boldsymbol a}\|^2 + \|\vec{\boldsymbol b}\|^2\right).$$

The squared lengths of the sides are

$$\|\overrightarrow{OA}\|^2 = \|\vec{\boldsymbol a}\|^2,\quad \|\overrightarrow{AB}\|^2 = \|\vec{\boldsymbol b}\|^2,\quad \|\overrightarrow{BC}\|^2 = \|\vec{\boldsymbol a}\|^2,\quad \|\overrightarrow{OC}\|^2 = \|\vec{\boldsymbol b}\|^2.$$

Together these also add up to $2\left(\|\vec{\boldsymbol a}\|^2 + \|\vec{\boldsymbol b}\|^2\right)$. $\square$

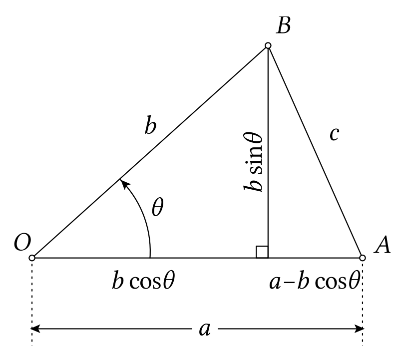


Figure 6. Proof of the law of cosines

5.6. The dot product and the angle between two vectors. Here is the most important interpretation of the dot product:

5.7. Theorem. *If the angle between two vectors $\vec{\boldsymbol a}$ and $\vec{\boldsymbol b}$ is θ, then one has*

$$\vec{\boldsymbol a}\bullet\vec{\boldsymbol b} = \|\vec{\boldsymbol a}\|\,\|\vec{\boldsymbol b}\|\,\cos\theta.$$

An important special case is where the vectors $\vec{\boldsymbol a}$ and $\vec{\boldsymbol b}$ are perpendicular. In that case $\theta = \frac{\pi}{2}$, so that $\cos\theta = 0$ and the dot product of $\vec{\boldsymbol a}$ and $\vec{\boldsymbol b}$ vanishes. Conversely, if both $\vec{\boldsymbol a}$ and $\vec{\boldsymbol b}$ are non zero vectors whose dot product vanishes, then $\cos\theta$ must vanish, and therefore $\theta = \frac{\pi}{2}$. In short, *two non-zero vectors are perpendicular if and only if their dot product vanishes.*

PROOF. We need *the law of cosines* from high-school trigonometry. Recall that for a triangle OAB with angle θ at the point O, and with sides OA and OB of lengths a and b, the length c of the opposing side AB is given by

$$c^2 = a^2 + b^2 - 2ab\cos\theta. \tag{75}$$

In trigonometry this is proved by dropping a perpendicular line from B onto the side OA. The triangle OAB gets divided into two right triangles, one of which has AB as hypotenuse. Pythagoras then implies

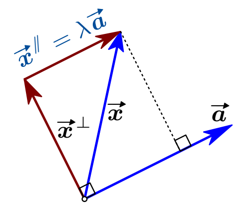
$$c^2 = (b\sin\theta)^2 + (a - b\cos\theta)^2.$$

After simplification you get (75).

To prove the theorem you let O be the origin, and then observe that the length of the side AB is the length of the vector $\overrightarrow{AB} = \vec{b} - \vec{a}$. Here $\vec{a} = \overrightarrow{OA}$, $\vec{b} = \overrightarrow{OB}$, and hence

$$c^2 = \|\vec{b} - \vec{a}\|^2 = (\vec{b} - \vec{a}) \bullet (\vec{b} - \vec{a}) = \|\vec{b}\|^2 + \|\vec{a}\|^2 - 2\vec{a} \bullet \vec{b}.$$

Compare this with (75), keeping in mind that $a = \|\vec{a}\|$ and $b = \|\vec{b}\|$: you are led to conclude that $-2\vec{a} \bullet \vec{b} = -2ab\cos\theta$, and thus $\vec{a} \bullet \vec{b} = \|\vec{a}\|\,\|\vec{b}\|\cos\theta$. $\square$



Given $\vec{x}$ and $\vec{a}$, find $\vec{x}^{\perp}$ and $\vec{x}^{/\!/}$.

5.8. Orthogonal projection of one vector onto another. The following construction comes up very often. Let $\vec{a} \neq \vec{0}$ be a given vector. Then for any other vector $\vec{x}$ there is a number λ such that

$$\vec{x} = \lambda\vec{a} + \vec{y}$$

where $\vec{y} \perp \vec{a}$. In other words, you can write any vector $\vec{x}$ as the sum of one vector parallel to $\vec{a}$ and another vector orthogonal to $\vec{a}$. The two vectors $\lambda\vec{a}$ and $\vec{y}$ are called the ***parallel*** and ***orthogonal components*** of the vector $\vec{x}$ (with respect to $\vec{a}$), and sometimes the following notation is used

$$\vec{x}^{/\!/} = \lambda\vec{a}, \qquad \vec{x}^{\perp} = \vec{y},$$

so that

$$\vec{x} = \vec{x}^{/\!/} + \vec{x}^{\perp}.$$

There are moderately simple formulas for $\vec{x}^{/\!/}$ and $\vec{x}^{\perp}$, but it is better to remember the following derivation of these formulas.

Assume that the vectors $\vec{a}$ and $\vec{x}$ are given. Then we look for a number λ such that $\vec{y} = \vec{x} - \lambda\vec{a}$ is perpendicular to $\vec{a}$. Recall that $\vec{a} \perp (\vec{x} - \lambda\vec{a})$ if and only if

$$\vec{a} \bullet (\vec{x} - \lambda\vec{a}) = 0.$$

Expand the dot product and you get this equation for λ

$$\vec{a} \bullet \vec{x} - \lambda\vec{a} \bullet \vec{a} = 0,$$

whence

$$\lambda = \frac{\vec{a} \bullet \vec{x}}{\vec{a} \bullet \vec{a}} = \frac{\vec{a} \bullet \vec{x}}{\|\vec{a}\|^2} \tag{76}$$

To compute the parallel and orthogonal components of $\vec{x}$ w.r.t. $\vec{a}$ you first compute λ according to (76), which tells you that the parallel component is given by

$$\vec{x}^{/\!/} = \lambda\vec{a} = \frac{\vec{a} \bullet \vec{x}}{\vec{a} \bullet \vec{a}}\,\vec{a}.$$

The orthogonal component is then "the rest," i.e. by definition $\vec{x}^{\perp} = \vec{x} - \vec{x}^{/\!/}$, so

$$\vec{x}^{\perp} = \vec{x} - \vec{x}^{/\!/} = \vec{x} - \frac{\vec{a} \bullet \vec{x}}{\vec{a} \bullet \vec{a}}\,\vec{a}.$$

5.9. Defining equations of lines. In § 3 we saw how to generate points on a line given two points on that line by means of a "parametrization." I.e. given points A and B on the line ℓ the point whose position vector is $\vec{x} = \vec{a} + t(\vec{b} - \vec{a})$ will be on ℓ for any value of the "parameter" t.

In this section we will use the dot-product to give a different description of lines in the plane (and planes in three dimensional space.) We will derive an equation for a line. Rather than generating points on the line ℓ this equation tells us if any given point X in the plane is on the line or not.

Here is the derivation of the equation of a line in the plane. To produce the equation you need two ingredients:

1. One particular point on the line (let's call this point A, and write $\vec{a}$ for its position vector),

2. a ***normal vector*** $\vec{n}$ for the line, i.e. a nonzero vector which is perpendicular to the line.

Now let X be any point in the plane, and consider the line segment AX.

- Clearly, X will be on the line if and only if AX is parallel to ℓ [1]
- Since ℓ is perpendicular to $\vec{n}$, the segment AX and the line ℓ will be parallel if and only if $AX \perp \vec{n}$.
- $AX \perp \vec{n}$ holds if and only if $\overrightarrow{AX} \cdot \vec{n} = 0$.

So in the end we see that X lies on the line ℓ if and only if the following vector equation is satisfied:

$$\overrightarrow{AX} \cdot \vec{n} = 0 \quad \text{or} \quad (\vec{x} - \vec{a}) \cdot \vec{n} = 0 \tag{77}$$

This equation is called a ***defining equation for the line*** ℓ.

Any given line has many defining equations. Just by changing the length of the normal you get a different equation, which still describes the same line.

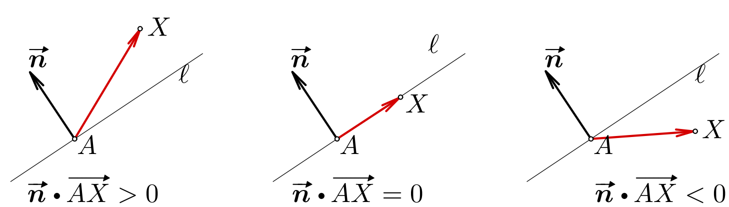


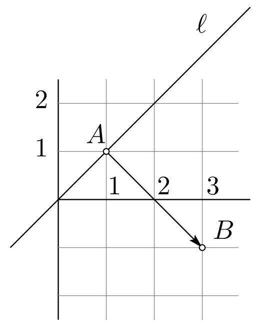
Figure 7. If $\vec{n}$ is a normal vector to the line ℓ, and if A is any point on ℓ, then there is a simple test which tells you if a point X is on ℓ or not: X is on ℓ if and only if $\overrightarrow{AX} \perp \vec{n}$, i.e. iff $\overrightarrow{AX} \cdot \vec{n} = 0$.

5.10. Line through one point and perpendicular to another line. *Find a defining equation for the line ℓ which goes through $A(1,1)$ and is perpendicular to the line segment AB where B is the point $(3,-1)$.*

Solution. We already know a point on the line, namely A, but we still need a normal vector. The line is required to be perpendicular to AB, so $\vec{n} = \overrightarrow{AB}$ is a normal vector:

$$\vec{n} = \overrightarrow{AB} = \begin{pmatrix} 3-1 \\ (-1)-1 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \end{pmatrix}$$

[1] From plane Euclidean geometry: parallel lines either don't intersect or they coincide.



Of course any multiple of $\vec{n}$ is also a normal vector, for instance

$$\vec{m} = \tfrac{1}{2}\vec{n} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

is a normal vector.

With $\vec{a} = \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ we then get the following equation for ℓ

$$\vec{n} \bullet (\vec{x} - \vec{a}) = \begin{pmatrix} 2 \\ -2 \end{pmatrix} \bullet \begin{pmatrix} x_1 - 1 \\ x_2 - 1 \end{pmatrix} = 2x_1 - 2x_2 = 0.$$

If you choose the normal $\vec{m}$ instead, you get

$$\vec{m} \bullet (\vec{x} - \vec{a}) = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \bullet \begin{pmatrix} x_1 - 1 \\ x_2 - 1 \end{pmatrix} = x_1 - x_2 = 0.$$

Both equations $2x_1 - 2x_2 = 0$ and $x_1 - x_2 = 0$ are equivalent and they both give defining equations for the line ℓ.

5.11. Distance to a line. Let ℓ be a line in the plane and assume a point A on the line as well as a vector $\vec{n}$ perpendicular to ℓ are known. Using the dot product one can easily compute the distance from the line to any other given point P in the plane. Here is how:

Draw the line m through A perpendicular to ℓ, and drop a perpendicular line from P onto m. let Q be the projection of P onto m. The distance from P to ℓ is then equal to the length of the line segment AQ. Since AQP is a right triangle one has

$$AQ = AP \cos\theta.$$

Here θ is the angle between the normal $\vec{n}$ and the vector $\overrightarrow{AP}$. One also has

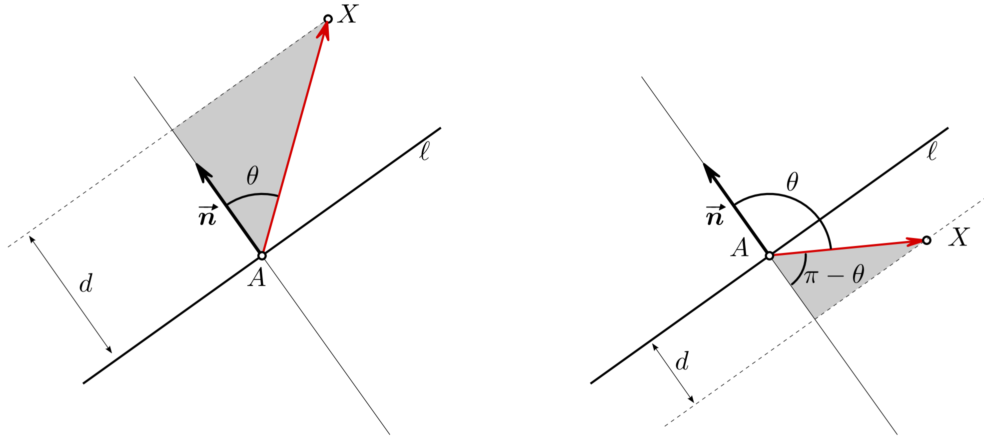
$$\vec{n} \bullet (\vec{p} - \vec{a}) = \vec{n} \bullet \overrightarrow{AP} = \|\overrightarrow{AP}\| \; \|\vec{n}\| \cos\theta = AP\, \|\vec{n}\| \cos\theta.$$

Hence we get

$$\mathrm{dist}(P, \ell) = \frac{\vec{n} \bullet (\vec{p} - \vec{a})}{\|\vec{n}\|}.$$

This argument from a drawing contains a hidden assumption, namely that the point P lies on the side of the line ℓ pointed to by the vector $\vec{n}$. If this is not the case, so that $\vec{n}$ and $\overrightarrow{AP}$ point to opposite sides of ℓ, then the angle between them exceeds $90°$, i.e. $\theta > \pi/2$. In this case $\cos\theta < 0$, and one has $AQ = -AP\cos\theta$. The distance formula therefore has to be modified to

$$\mathrm{dist}(P, \ell) = -\frac{\vec{n} \bullet (\vec{p} - \vec{a})}{\|\vec{n}\|}.$$



$$\vec{n} \cdot \overrightarrow{AX} > 0 \qquad d = \|\overrightarrow{AX}\| \cos\theta$$

$$\vec{n} \cdot \overrightarrow{AX} < 0 \qquad \begin{aligned} d &= \|\overrightarrow{AX}\| \cos(\pi - \theta) \\ &= -\|\overrightarrow{AX}\| \cos\theta \end{aligned}$$

Figure 8. The distance from a point P to a line ℓ.

We do not need to know in advance which formula to use. If we compute $\vec{n} \cdot (\vec{p} - \vec{a})$ and find that it is negative then we know that the normal vector and the point are on opposite sides of the line ℓ. In either case the distance is given by

$$\text{dist}(P, \ell) = \left| \frac{\vec{n} \cdot (\vec{p} - \vec{a})}{\|\vec{n}\|} \right|.$$

5.12. Defining equation of a plane. Just as we have seen how we can form the defining equation for a line in the plane from just one point on the line and one normal vector to the line, we can also form the defining equation for a plane in space, again knowing only one point on the plane, and a vector perpendicular to it. If A is a point on some plane $\mathcal{P}$ and $\vec{n}$ is a vector perpendicular to $\mathcal{P}$, then any other point X lies on $\mathcal{P}$ if and only if $\overrightarrow{AX} \perp \vec{n}$. In other words, in terms of the position vectors $\vec{a}$ and $\vec{x}$ of A and X,

$$\text{the point } X \text{ is on } \mathcal{P} \iff \vec{n} \cdot (\vec{x} - \vec{a}) = 0.$$

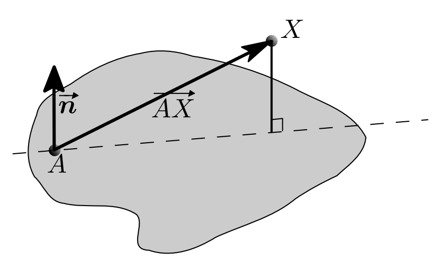


Figure 9. A point P lies on the plane if the vector $\overrightarrow{AP}$ is perpendicular to $\vec{n}$.

Arguing just as in § 5.11 you find that the distance of a point X in space to the plane $\mathcal{P}$ is

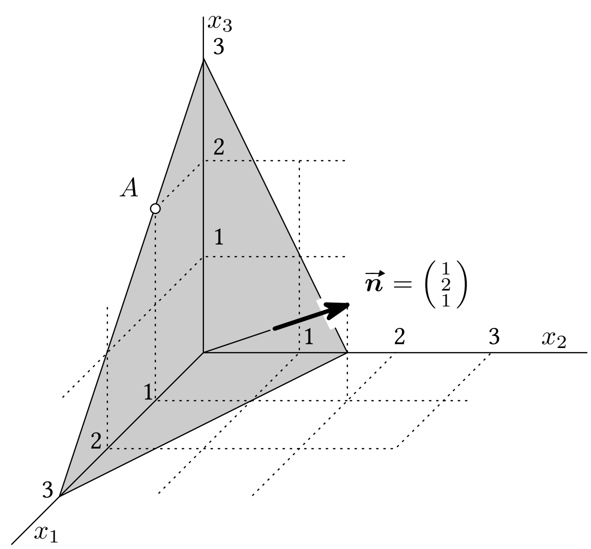
$$\operatorname{dist}(X, \mathcal{P}) = \pm \frac{\vec{n} \bullet (\vec{x} - \vec{a})}{\|\vec{n}\|}. \tag{78}$$

Here the sign is "$+$" if X and the normal $\vec{n}$ are on the same side of the plane $\mathcal{P}$; otherwise the sign is "$-$".

5.13. Example. Find the defining equation for the plane $\mathcal{P}$ through the point $A(1,0,2)$ which is perpendicular to the vector $\begin{pmatrix}1\\2\\1\end{pmatrix}$.

Solution: We know a point (A) and a normal vector $\vec{n} = \begin{pmatrix}1\\2\\1\end{pmatrix}$ for $\mathcal{P}$. Then any point X with coordinates (x_1, x_2, x_3), or, with position vector $\vec{x} = \begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix}$, will lie on the plane $\mathcal{P}$ if and only if

$$\begin{aligned}
\vec{n} \bullet (\vec{x} - \vec{a}) = 0 \iff & \begin{pmatrix}1\\2\\1\end{pmatrix} \bullet \left\{ \begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} - \begin{pmatrix}1\\0\\2\end{pmatrix} \right\} = 0 \\
\iff & \begin{pmatrix}1\\2\\1\end{pmatrix} \bullet \begin{pmatrix}x_1 - 1\\x_2\\x_3 - 2\end{pmatrix} = 0 \\
\iff & 1 \cdot (x_1 - 1) + 2 \cdot (x_2) + 1 \cdot (x_3 - 2) = 0 \\
\iff & x_1 + 2x_2 + x_3 - 3 = 0.
\end{aligned}$$



5.14. Example continued. Let $\mathcal{P}$ be the plane from the previous example. Which of the points $P(0,0,1)$, $Q(0,0,2)$, $R(-1,2,0)$ and $S(-1,0,5)$ lie on $\mathcal{P}$? Compute the distances from the points P, Q, R, S to the plane $\mathcal{P}$. Separate the points which do not lie on $\mathcal{P}$ into two group of points which lie on the same side of $\mathcal{P}$.

Solution: We apply (78) to the position vectors $\vec{\boldsymbol{p}}, \vec{\boldsymbol{q}}, \vec{\boldsymbol{r}}, \vec{\boldsymbol{s}}$ of the points P, Q, R, S. For each calculation we need

$$\|\vec{\boldsymbol{n}}\| = \sqrt{1^2+2^2+1^2} = \sqrt{6}.$$

The third component of the given normal $\vec{\boldsymbol{n}} = \left(\begin{smallmatrix}1\\2\\1\end{smallmatrix}\right)$ is positive, so $\vec{\boldsymbol{n}}$ points "upwards." Therefore, if a point lies on the side of $\mathcal{P}$ pointed to by $\vec{\boldsymbol{n}}$, we shall say that the point lies *above the plane.*

P: $\vec{\boldsymbol{p}} = \left(\begin{smallmatrix}0\\0\\1\end{smallmatrix}\right)$, $\vec{\boldsymbol{p}} - \vec{\boldsymbol{a}} = \left(\begin{smallmatrix}-1\\0\\-1\end{smallmatrix}\right)$, $\vec{\boldsymbol{n}} \bullet (\vec{\boldsymbol{p}} - \vec{\boldsymbol{a}}) = 1\cdot(-1) + 2\cdot(0) + 1\cdot(-1) = -2$

$$\frac{\vec{\boldsymbol{n}} \bullet (\vec{\boldsymbol{p}} - \vec{\boldsymbol{a}})}{\|\vec{\boldsymbol{n}}\|} = -\frac{2}{\sqrt{6}} = -\frac{1}{3}\sqrt{6}.$$

This quantity is negative, so P lies below $\mathcal{P}$. Its distance to $\mathcal{P}$ is $\frac{1}{3}\sqrt{6}$.

Q: $\vec{\boldsymbol{q}} = \left(\begin{smallmatrix}0\\0\\2\end{smallmatrix}\right)$, $\vec{\boldsymbol{p}} - \vec{\boldsymbol{a}} = \left(\begin{smallmatrix}-1\\0\\0\end{smallmatrix}\right)$, $\vec{\boldsymbol{n}} \bullet (\vec{\boldsymbol{p}} - \vec{\boldsymbol{a}}) = 1\cdot(-1) + 2\cdot(0) + 1\cdot(0) = -1$

$$\frac{\vec{\boldsymbol{n}} \bullet (\vec{\boldsymbol{p}} - \vec{\boldsymbol{a}})}{\|\vec{\boldsymbol{n}}\|} = -\frac{1}{\sqrt{6}} = -\frac{1}{6}\sqrt{6}.$$

This quantity is negative, so Q also lies below $\mathcal{P}$. Its distance to $\mathcal{P}$ is $\frac{1}{6}\sqrt{6}$.

R: $\vec{\boldsymbol{r}} = \left(\begin{smallmatrix}-1\\2\\0\end{smallmatrix}\right)$, $\vec{\boldsymbol{p}} - \vec{\boldsymbol{a}} = \left(\begin{smallmatrix}-2\\2\\-2\end{smallmatrix}\right)$, $\vec{\boldsymbol{n}} \bullet (\vec{\boldsymbol{p}} - \vec{\boldsymbol{a}}) = 1\cdot(-2) + 2\cdot(2) + 1\cdot(-2) = 0$

$$\frac{\vec{\boldsymbol{n}} \bullet (\vec{\boldsymbol{p}} - \vec{\boldsymbol{a}})}{\|\vec{\boldsymbol{n}}\|} = 0.$$

Thus R lies on the plane $\mathcal{P}$, and its distance to $\mathcal{P}$ is of course 0.

S: $\vec{\boldsymbol{s}} = \left(\begin{smallmatrix}-1\\0\\5\end{smallmatrix}\right)$, $\vec{\boldsymbol{p}} - \vec{\boldsymbol{a}} = \left(\begin{smallmatrix}-2\\0\\3\end{smallmatrix}\right)$, $\vec{\boldsymbol{n}} \bullet (\vec{\boldsymbol{p}} - \vec{\boldsymbol{a}}) = 1\cdot(-1) + 2\cdot(0) + 1\cdot(3) = 2$

$$\frac{\vec{\boldsymbol{n}} \bullet (\vec{\boldsymbol{p}} - \vec{\boldsymbol{a}})}{\|\vec{\boldsymbol{n}}\|} = \frac{2}{\sqrt{6}} = \frac{1}{3}\sqrt{6}.$$

This quantity is positive, so S lies above $\mathcal{P}$. Its distance to $\mathcal{P}$ is $\frac{1}{3}\sqrt{6}$.

We have found that P and Q lie below the plane, R lies on the plane, and S is above the plane.

5.15. Where does the line through the points $B(2,0,0)$ and $C(0,1,2)$ intersect the plane $\mathcal{P}$ from example 5.13? *Solution:* Let ℓ be the line through B and C. We set up the parametric equation for ℓ. According to §3, (69) every point X on ℓ has position vector $\vec{\boldsymbol{x}}$ given by

$$\vec{\boldsymbol{x}} = \vec{\boldsymbol{b}} + t(\vec{\boldsymbol{c}} - \vec{\boldsymbol{b}}) = \begin{pmatrix}2\\0\\0\end{pmatrix} + t\begin{pmatrix}0-2\\1-0\\2-0\end{pmatrix} = \begin{pmatrix}2-2t\\t\\2t\end{pmatrix} \tag{79}$$

for some value of t.

The point X whose position vector $\vec{\boldsymbol{x}}$ is given above lies on the plane $\mathcal{P}$ if $\vec{\boldsymbol{x}}$ satisfies the defining equation of the plane. In example 5.13 we found this defining equation. It was

$$\vec{\boldsymbol{n}} \bullet (\vec{\boldsymbol{x}} - \vec{\boldsymbol{a}}) = 0, \text{ i.e. } x_1 + 2x_2 + x_3 - 3 = 0. \tag{80}$$

So to find the point of intersection of ℓ and $\mathcal{P}$ you substitute the parametrization (79) in the defining equation (80):

$$0 = x_1 + 2x_2 + x_3 - 3 = (2-2t) + 2(t) + (2t) - 3 = 2t - 1.$$

This implies $t = \frac{1}{2}$, and thus the intersection point has position vector

$$\vec{x} = \vec{b} + \tfrac{1}{2}(\vec{c} - \vec{b}) = \begin{pmatrix} 2-2t \\ t \\ 2t \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{1}{2} \\ 1 \end{pmatrix},$$

i.e. ℓ and $\mathcal{P}$ intersect at $X(1, \frac{1}{2}, 1)$.

6. Cross Product

6.1. Algebraic definition of the cross product. Here is the definition of the cross-product of two vectors. The definition looks a bit strange and arbitrary at first sight – it really makes you wonder who thought of this. We will just put up with that for now and explore the properties of the cross product. Later on we will see a geometric interpretation of the cross product which will show that this particular definition is really useful. We will also find a few tricks that will help you reproduce the formula without memorizing it.

6.2. Definition. *The "outer product" or "cross product" of two vectors is given by*

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{pmatrix}$$

Note that the cross-product of two vectors is again a vector!

6.3. Example. If you set $\vec{b} = \vec{a}$ in the definition you find the following important fact: *The cross product of any vector with itself is the zero vector:*

$$\vec{a} \times \vec{a} = \vec{0} \quad \text{for } \textbf{\textit{any}} \text{ vector } \vec{a}.$$

6.4. Example. Let $\vec{a} = \left(\begin{smallmatrix} 1 \\ 2 \\ 3 \end{smallmatrix}\right)$, $\vec{b} = \left(\begin{smallmatrix} -2 \\ 1 \\ 0 \end{smallmatrix}\right)$ and compute the cross product of these vectors.

Solution:

$$\vec{a} \times \vec{b} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \times \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2\cdot 0 - 3\cdot 1 \\ 3\cdot(-2) - 1\cdot 0 \\ 1\cdot 1 - 2\cdot(-2) \end{pmatrix} = \begin{pmatrix} -3 \\ -6 \\ 5 \end{pmatrix}$$

6.5. Algebraic properties of the cross product. Unlike the dot product, the cross product of two vectors behaves much less like ordinary multiplication. To begin with, the product is ***not commutative*** – instead one has

$$\vec{a} \times \vec{b} = -\vec{b} \times \vec{a} \quad \text{for } \textbf{\textit{all}} \text{ vectors } \vec{a} \text{ and } \vec{b}. \tag{81}$$

This property is sometimes called "anti-commutative."

Since the crossproduct of two vectors is again a vector you can compute the cross product of three vectors $\vec{a}, \vec{b}, \vec{c}$. You now have a choice: do you first multiply $\vec{a}$ and $\vec{b}$, or $\vec{b}$ and $\vec{c}$, or $\vec{a}$ and $\vec{c}$? With numbers it makes no difference (e.g. $2 \times (3 \times 5) = 2 \times 15 = 30$ and $(2 \times 3) \times 5 = 6 \times 5 =$ also 30) but with the cross product of vectors it does matter: the cross product is ***not associative,*** i.e.

$$\vec{a} \times (\vec{b} \times \vec{c}) \neq (\vec{a} \times \vec{b}) \times \vec{c} \quad \text{for } \textbf{\textit{most}} \text{ vectors } \vec{a}, \vec{b}, \vec{c}.$$

$\vec{\imath} \times (\vec{\imath} \times \vec{\jmath}) = -\vec{\jmath}$,
and
$(\vec{\imath} \times \vec{\imath}) \times \vec{\jmath} = \vec{0}$,
so
$\vec{\imath} \times (\vec{\imath} \times \vec{\jmath}) \neq (\vec{\imath} \times \vec{\imath}) \times \vec{\jmath}$
Conclusion:
"$\times$" is not associative

The ***distributive law*** does hold, i.e.

$$\vec{a} \times (\vec{b} + \vec{c}) = \vec{a} \times \vec{b} + \vec{a} \times \vec{c}, \quad \text{and} \quad (\vec{b} + \vec{c}) \times \vec{a} = \vec{b} \times \vec{a} + \vec{c} \times \vec{a}$$

is true for all vectors $\vec{a}$, $\vec{b}$, $\vec{c}$.

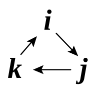
Also, an associative law, where one of the factors is a number and the other two are vectors, does hold. I.e.

$$t(\vec{a}\times\vec{b}) = (t\vec{a})\times\vec{b} = \vec{a}\times(t\vec{b})$$

holds for all vectors $\vec{a}$, $\vec{b}$ and any number t. We were already using these properties when we multiplied $(a_1\vec{\imath}+a_2\vec{\jmath}+a_3\vec{k})\times(b_1\vec{\imath}+b_2\vec{\jmath}+b_3\vec{k})$ in the previous section.

Finally, the cross product is only defined for space vectors, not for plane vectors.

$\times$	$\vec{\imath}$	$\vec{\jmath}$	$\vec{k}$
$\vec{\imath}$	$\vec{0}$	$\vec{k}$	$-\vec{\jmath}$
$\vec{\jmath}$	$-\vec{k}$	$\vec{0}$	$\vec{\imath}$
$\vec{k}$	$\vec{\jmath}$	$-\vec{\imath}$	$\vec{0}$



6.6. Ways to compute the cross product. In terms of the standard basis vectors you can check the *multiplication table*. An easy way to remember the multiplication table is to put the vectors $\vec{\imath}$, $\vec{\jmath}$, $\vec{k}$ clockwise in a circle. Given two of the three vectors their product is either plus or minus the remaining vector. To determine the sign you step from the first vector to the second, to the third: if this makes you go clockwise you have a plus sign, if you have to go counterclockwise, you get a minus.

The products of $\vec{\imath}$, $\vec{\jmath}$ and $\vec{k}$ are all you need to know to compute the cross product. Given two vectors $\vec{a}$ and $\vec{b}$ write them as $\vec{a} = a_1\vec{\imath}+a_2\vec{\jmath}+a_3\vec{k}$ and $\vec{b} = b_1\vec{\imath}+b_2\vec{\jmath}+b_3\vec{k}$, and multiply as follows

$$\begin{aligned}
\vec{a}\times\vec{b} &= (a_1\vec{\imath}+a_2\vec{\jmath}+a_3\vec{k})\times(b_1\vec{\imath}+b_2\vec{\jmath}+b_3\vec{k})\\
&= a_1\vec{\imath}\times(b_1\vec{\imath}+b_2\vec{\jmath}+b_3\vec{k})\\
&\quad +a_2\vec{\jmath}\times(b_1\vec{\imath}+b_2\vec{\jmath}+b_3\vec{k})\\
&\quad +a_3\vec{k}\times(b_1\vec{\imath}+b_2\vec{\jmath}+b_3\vec{k})\\
&= a_1b_1\vec{\imath}\times\vec{\imath} + a_1b_2\vec{\imath}\times\vec{\jmath} + a_1b_3\vec{\imath}\times\vec{k} +\\
&\quad a_2b_1\vec{\jmath}\times\vec{\imath} + a_2b_2\vec{\jmath}\times\vec{\jmath} + a_2b_3\vec{\jmath}\times\vec{k} +\\
&\quad a_3b_1\vec{k}\times\vec{\imath} + a_3b_2\vec{k}\times\vec{\jmath} + a_3b_3\vec{k}\times\vec{k}\\
&= a_1b_1\vec{0} + a_1b_2\vec{k} - a_1b_3\vec{\jmath}\\
&\quad -a_2b_1\vec{k} + a_2b_2\vec{0} + a_2b_3\vec{\imath} +\\
&\quad a_3b_1\vec{\jmath} - a_3b_2\vec{\imath} + a_3b_3\vec{0}\\
&= (a_2b_3-a_3b_2)\vec{\imath}+(a_3b_1-a_1b_3)\vec{\jmath}+(a_1b_2-a_2b_1)\vec{k}
\end{aligned}$$

This is a useful way of remembering how to compute the cross product, particularly when many of the components a_i and b_j are zero.

6.7. Example. Compute $\vec{k}\times(p\vec{\imath}+q\vec{\jmath}+r\vec{k})$:

$$\vec{k}\times(p\vec{\imath}+q\vec{\jmath}+r\vec{k}) = p(\vec{k}\times\vec{\imath})+q(\vec{k}\times\vec{\jmath})+r(\vec{k}\times\vec{k}) = -q\vec{\imath}+p\vec{\jmath}.$$

There is another way of remembering how to find $\vec{a}\times\vec{b}$. It involves the "triple product" and determinants. See § 6.8.

6.8. The triple product and determinants.

6.9. Definition. *The triple product of three given vectors $\vec{a}$, $\vec{b}$, and $\vec{c}$ is defined to be*

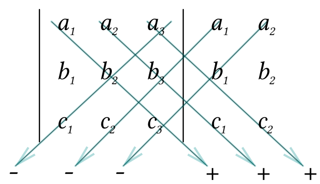
$$\vec{a}\cdot(\vec{b}\times\vec{c}).$$

In terms of the components of $\vec{a}$, $\vec{b}$, and $\vec{c}$ one has

$$\vec{a}\cdot(\vec{b}\times\vec{c}) = \begin{pmatrix}a_1\\a_2\\a_3\end{pmatrix}\cdot\begin{pmatrix}b_2c_3-b_3c_2\\b_3c_1-b_1c_3\\b_1c_2-b_2c_1\end{pmatrix}$$
$$= a_1b_2c_3 - a_1b_3c_2 + a_2b_3c_1 - a_2b_1c_3 + a_3b_1c_2 - a_3b_2c_1.$$

This quantity is called a ***determinant,*** and is written as follows

$$\begin{vmatrix}a_1 & b_1 & c_1\\ a_2 & b_2 & c_2\\ a_3 & b_3 & c_3\end{vmatrix} = a_1b_2c_3 - a_1b_3c_2 + a_2b_3c_1 - a_2b_1c_3 + a_3b_1c_2 - a_3b_2c_1 \tag{82}$$



There's a useful shortcut for computing such a determinant: after writing the determinant, append a fourth and a fifth column which are just copies of the first two columns of the determinant. The determinant then is the sum of six products, one for each dotted line in the drawing. Each term has a sign: if the factors are read from top-left to bottom-right, the term is positive, if they are read from top-right to bottom left the term is negative.
This shortcut is also very useful for computing the crossproduct. To compute the cross product of two given vectors $\vec{a}$ and $\vec{b}$ you arrange their components in the following determinant

$$\vec{a}\times\vec{b} = \begin{vmatrix}\vec{\imath} & a_1 & b_1\\ \vec{\jmath} & a_2 & b_2\\ \vec{k} & a_3 & b_3\end{vmatrix} = (a_2b_3-a_3b_2)\vec{\imath} + (a_3b_1-a_1b_3)\vec{\jmath} + (a_1b_2-a_2b_1)\vec{k}. \tag{83}$$

This is not a normal determinant since some of its entries are vectors, but if you ignore that odd circumstance and simply compute the determinant according to the definition (82), you get (83).

An important property of the triple product is that it is much more symmetric in the factors $\vec{a}, \vec{b}, \vec{c}$ than the notation $\vec{a}\cdot(\vec{b}\times\vec{c})$ suggests.

6.10. Theorem. *For any triple of vectors $\vec{a}, \vec{b}, \vec{c}$ one has*

$$\vec{a}\cdot(\vec{b}\times\vec{c}) = \vec{b}\cdot(\vec{c}\times\vec{a}) = \vec{c}\cdot(\vec{a}\times\vec{b}),$$

and

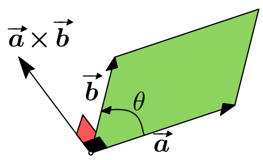
$$\vec{a}\cdot(\vec{b}\times\vec{c}) = -\vec{b}\cdot(\vec{a}\times\vec{c}) = -\vec{c}\cdot(\vec{b}\times\vec{a}).$$

In other words, if you exchange two factors in the product $\vec{a}\cdot(\vec{b}\times\vec{c})$ it changes its sign. If you "rotate the factors," i.e. if you replace $\vec{a}$ by $\vec{b}$, $\vec{b}$ by $\vec{c}$ and $\vec{c}$ by $\vec{a}$, the product doesn't change at all.

6.11. Geometric description of the cross product.

6.12. Theorem.

$$\vec{a}\times\vec{b} \perp \vec{a}, \vec{b}$$



PROOF. We use the triple product:

$$\vec{a}\cdot(\vec{a}\times\vec{b}) = \vec{b}\cdot(\vec{a}\times\vec{a}) = 0$$

since $\vec{a}\times\vec{a} = \vec{0}$ for any vector $\vec{a}$. It follows that $\vec{a}\times\vec{b}$ is perpendicular to $\vec{a}$.

Similarly, $\vec{b}\cdot(\vec{a}\times\vec{b}) = \vec{a}\cdot(\vec{b}\times\vec{b}) = 0$ shows that $\vec{a}\cdot\vec{b}$ is perpendicular to $\vec{b}$. □

6.13. Theorem.

$$\|\vec{a}\times\vec{b}\| = \|\vec{a}\|\,\|\vec{b}\|\,\sin\theta$$

PROOF. Bruce[2] just slipped us a piece of paper with the following formula on it:

$$\|\vec{a}\times\vec{b}\|^2 + (\vec{a}\cdot\vec{b})^2 = \|\vec{a}\|^2\|\vec{b}\|^2. \tag{84}$$

After setting $\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ and $\vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$ and diligently computing both sides we find that this formula actually holds for any pair of vectors $\vec{a}, \vec{b}$! The (long) computation which implies this identity will be presented in class (maybe).

If we assume that Lagrange's identity holds then we get

$$\|\vec{a}\times\vec{b}\|^2 = \|\vec{a}\|^2\|\vec{b}\|^2-(\vec{a}\cdot\vec{b})^2 = \|\vec{a}\|^2\|\vec{b}\|^2-\|\vec{a}\|^2\|\vec{b}\|^2\cos^2\theta = \|\vec{a}\|^2\|\vec{b}\|^2\sin^2\theta$$

since $1-\cos^2\theta = \sin^2\theta$. The theorem is proved. □

These two theorems *almost* allow you to construct the cross product of two vectors geometrically. If $\vec{a}$ and $\vec{b}$ are two vectors, then their cross product satisfies the following description:

(1) If $\vec{a}$ and $\vec{b}$ are parallel, then the angle θ between them vanishes, and so their cross product is the zero vector. Assume from here on that $\vec{a}$ and $\vec{b}$ are not parallel.
(2) $\vec{a}\times\vec{b}$ is perpendicular to both $\vec{a}$ and $\vec{b}$. In other words, since $\vec{a}$ and $\vec{b}$ are not parallel, they determine a plane, and their cross product is a vector perpendicular to this plane.
(3) the length of the cross product $\vec{a}\times\vec{b}$ is $\|\vec{a}\|\;\|\vec{b}\|\,\sin\theta$.

There are only two vectors that satisfy conditions 2 and 3: to determine which one of these is the cross product you must apply the ***Right Hand Rule*** (screwdriver rule, corkscrew rule, etc.) for $\vec{a}, \vec{b}, \vec{a}\times\vec{b}$: if you turn a screw whose axis is perpendicular to $\vec{a}$ and $\vec{b}$ in the direction from $\vec{a}$ to $\vec{b}$, the screw moves in the direction of $\vec{a}\times\vec{b}$.

Alternatively, without seriously injuring yourself, you should be able to make a fist with your ***right*** hand, and then stick out your thumb, index and middle fingers so that your thumb is $\vec{a}$, your index finger is $\vec{b}$ and your middle finger is $\vec{a}\times\vec{b}$. If you do this with your left hand you will get $-\vec{a}\times\vec{b}$ instead of $\vec{a}\times\vec{b}$.

[2]It's actually called ***Lagrange's identity***. Yes, the same Lagrange who found the formula for the remainder term in Taylor's formula.

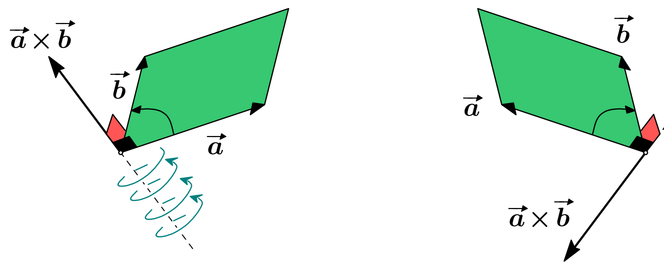


Figure 10. The right hand rule for the cross product.

7. A few applications of the cross product

7.1. Area of a parallelogram. Let $ABCD$ be a parallelogram. Its area is given by "height times base," a formula which should be familiar from high school geometry.

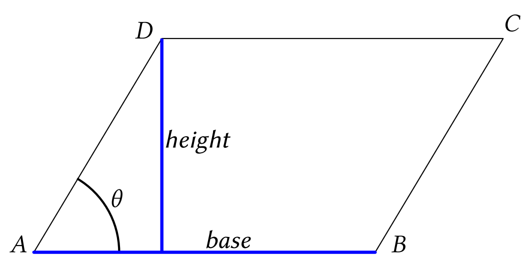


Figure 11. The area of a parallelogram

If the angle between the sides AB and AD is θ, then the height of the parallelogram is $\|\overrightarrow{AD}\|\sin\theta$, so that the area of $ABCD$ is

$$\text{area of } ABCD = \|\overrightarrow{AB}\|\;\|\overrightarrow{AD}\|\sin\theta = \|\overrightarrow{AB}\times\overrightarrow{AD}\|\,. \tag{85}$$

The area of the triangle ABD is of course half as much,

$$\text{area of triangle } ABD = \tfrac{1}{2}\,\|\overrightarrow{AB}\times\overrightarrow{AD}\|\,.$$

These formulae are valid even when the points A, B, C, and D are points in space. Of course they must lie in one plane for otherwise $ABCD$ couldn't be a parallelogram.

7.2. Example. Let the points $A(1,0,2)$, $B(2,0,0)$, $C(3,1,-1)$ and $D(2,1,1)$ be given.

Show that $ABCD$ is a parallelogram, and compute its area.

Solution: $ABCD$ will be a parallelogram if and only if $\overrightarrow{AC} = \overrightarrow{AB} + \overrightarrow{AD}$. In terms of the position vectors $\vec{a}, \vec{b}$, $\vec{c}$ and $\vec{d}$ of A, B, C, D this boils down to

$$\vec{c} - \vec{a} = (\vec{b} - \vec{a}) + (\vec{d} - \vec{a}), \quad \text{i.e.} \quad \vec{a} + \vec{c} = \vec{b} + \vec{d}.$$

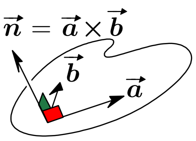
For our points we get

$$\vec{a} + \vec{c} = \begin{pmatrix}1\\0\\2\end{pmatrix} + \begin{pmatrix}3\\1\\-1\end{pmatrix} = \begin{pmatrix}4\\1\\1\end{pmatrix}, \qquad \vec{b} + \vec{d} = \begin{pmatrix}2\\0\\0\end{pmatrix} + \begin{pmatrix}2\\1\\1\end{pmatrix} = \begin{pmatrix}4\\1\\1\end{pmatrix}.$$

So $ABCD$ is indeed a parallelogram. Its area is the length of

$$\overrightarrow{AB}\times\overrightarrow{AD} = \begin{pmatrix} 2-1 \\ 0 \\ 0-2 \end{pmatrix} \times \begin{pmatrix} 2-1 \\ 1-0 \\ 1-2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} \times \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} -2 \\ -1 \\ -1 \end{pmatrix}.$$

So the area of $ABCD$ is $\sqrt{(-2)^2+(-1)^2+(-1)^2} = \sqrt{6}$.



7.3. Finding the normal to a plane. If you know two vectors $\vec{a}$ and $\vec{b}$ which are parallel to a given plane $\mathcal{P}$ but not parallel to each other, then you can find a normal vector for the plane $\mathcal{P}$ by computing

$$\vec{n} = \vec{a}\times\vec{b}.$$

We have just seen that the vector $\vec{n}$ must be perpendicular to both $\vec{a}$ and $\vec{b}$, and hence[3] it is perpendicular to the plane $\mathcal{P}$.

This trick is especially useful when you have three points A, B and C, and you want to find the defining equation for the plane $\mathcal{P}$ through these points. We will assume that the three points do not all lie on one line, for otherwise there are many planes through A, B and C.

To find the defining equation we need one point on the plane (we have three of them), and a normal vector to the plane. A normal vector can be obtained by computing the cross product of two vectors parallel to the plane. Since $\overrightarrow{AB}$ and $\overrightarrow{AC}$ are both parallel to $\mathcal{P}$, the vector $\vec{n} = \overrightarrow{AB}\times\overrightarrow{AC}$ is such a normal vector.

Thus the defining equation for the plane through three given points A, B and C is

$$\vec{n}\bullet(\vec{x}-\vec{a}) = 0, \quad \text{with} \quad \vec{n} = \overrightarrow{AB}\times\overrightarrow{AC} = (\vec{b}-\vec{a})\times(\vec{c}-\vec{a}).$$

7.4. Example. Find the defining equation of the plane $\mathcal{P}$ through the points $A(2,-1,0)$, $B(2,1,-1)$ and $C(-1,1,1)$. Find the intersections of $\mathcal{P}$ with the three coordinate axes, and find the distance from the origin to $\mathcal{P}$.

Solution: We have

$$\overrightarrow{AB} = \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} \quad \text{and} \quad \overrightarrow{AC} = \begin{pmatrix} -3 \\ 2 \\ 1 \end{pmatrix}$$

so that

$$\vec{n} = \overrightarrow{AB}\times\overrightarrow{AC} = \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} \times \begin{pmatrix} -3 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \\ 6 \end{pmatrix}$$

is a normal to the plane. The defining equation for $\mathcal{P}$ is therefore

$$0 = \vec{n}\bullet(\vec{x}-\vec{a}) = \begin{pmatrix} 4 \\ 3 \\ 6 \end{pmatrix} \bullet \begin{pmatrix} x_1-2 \\ x_2+1 \\ x_3-0 \end{pmatrix}$$

i.e.

$$4x_1 + 3x_2 + 6x_3 - 5 = 0.$$

The plane intersects the x_1 axis when $x_2 = x_3 = 0$ and hence $4x_1 - 5 = 0$, i.e. in the point $(\frac{5}{4}, 0, 0)$. The intersections with the other two axes are $(0, \frac{5}{3}, 0)$ and $(0, 0, \frac{5}{6})$.

The distance from any point with position vector $\vec{x}$ to $\mathcal{P}$ is given by

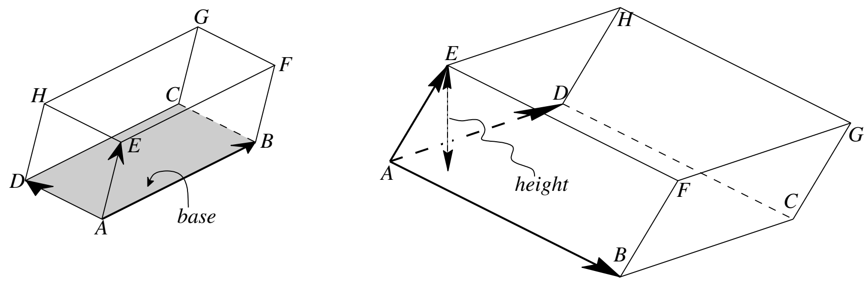
$$\text{dist} = \pm\frac{\vec{n}\bullet(\vec{x}-\vec{a})}{\|\vec{n}\|},$$

[3]This statement needs a proof which we will skip. Instead have a look at the picture

so the distance from the origin (whose position vector is $\vec{x} = \vec{0} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$) to $\mathcal{P}$ is

$$\text{distance origin to } \mathcal{P} = \pm \frac{\vec{a} \bullet \vec{n}}{\|\vec{n}\|} = \pm \frac{2 \cdot 4 + (-1) \cdot 3 + 0 \cdot 6}{\sqrt{4^2 + 3^2 + 6^2}} = \frac{5}{\sqrt{61}} (\approx 1.024 \cdots).$$

7.5. Volume of a parallelepiped.



A ***parallelepiped*** is a three dimensional body whose sides are parallelograms. For instance, a cube is an example of a parallelepiped; a rectangular block (whose faces are rectangles, meeting at right angles) is also a parallelepiped. Any parallelepiped has 8 vertices (corner points), 12 edges and 6 faces.

Let $\genfrac{}{}{0pt}{}{ABCD}{EFGH}$ be a parallelepiped. If we call one of the faces, say $ABCD$, the base of the parallelepiped, then the other face $EFGH$ is parallel to the base. The ***height of the parallelepiped*** is the distance from any point in $EFGH$ to the base, e.g. to compute the height of $\genfrac{}{}{0pt}{}{ABCD}{EFGH}$ one could compute the distance from the point E (or F, or G, or H) to the plane through $ABCD$.

The volume of the parallelepiped $\genfrac{}{}{0pt}{}{ABCD}{EFGH}$ is given by the formula

$$\text{Volume } \genfrac{}{}{0pt}{}{ABCD}{EFGH} = \text{Area of base} \times \text{height}.$$

Since the base is a parallelogram we know its area is given by

$$\text{Area of base} ABCD = \|\overrightarrow{AB} \times \overrightarrow{AD}\|$$

We also know that $\vec{n} = \overrightarrow{AB} \times \overrightarrow{AD}$ is a vector perpendicular to the plane through $ABCD$, i.e. perpendicular to the base of the parallelepiped. If we let the angle between the edge AE and the normal $\vec{n}$ be ψ, then the height of the parallelepiped is given by

$$\text{height} = \|\overrightarrow{AE}\| \cos \psi.$$

Therefore the triple product of $\overrightarrow{AB}$, $\overrightarrow{AD}$, $\overrightarrow{AE}$ is

$$\begin{aligned} \text{Volume } \genfrac{}{}{0pt}{}{ABCD}{EFGH} &= \text{height} \times \text{Area of base} \\ &= \|\overrightarrow{AE}\| \cos \psi \quad \|\overrightarrow{AB} \times \overrightarrow{AD}\|, \end{aligned}$$

i.e.

$$\boxed{\text{Volume } \genfrac{}{}{0pt}{}{ABCD}{EFGH} = \overrightarrow{AE} \bullet (\overrightarrow{AB} \times \overrightarrow{AD}).}$$

8. Notation

When we use vectors to describe the real world it is very important to distinguish between a point A, its position vector $\vec{a} = \overrightarrow{OA}$, and its coordinates a_1, a_2, a_3.

- A ***point*** can be any point in space. For instance, A could be the top of the Lincoln statue on Bascom Hill. That is a well-defined location.

- To specify the ***position vector*** for a point we must also specify which point we decide to call the origin (the center of the Capitol; or the center of the Earth; or the center of the Sun; etc.). Different choices of origins give us different position vectors for the same point.
- To specify the ***coordinates*** of the point A we not only need an origin, but we also have to specify three perpendicular axes, which we call the x, y, and z axes (or we can give them different names).

Given a point in the plane, or in space we can form its position vector. So associated to a point we have three different mathematical objects: the point, its position vector and its coordinates. Table 1 summarises these different notations.

Common abuses of notation that should be avoided. Sometimes students will write things like

$$\vec{a} = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix} = 6, \qquad \textit{(aargh!)}$$

which really can't be true because $\left(\begin{smallmatrix} 3 \\ 1 \\ 2 \end{smallmatrix}\right)$ is a vector and 6 is a number. ***Vectors and numbers are not equal!***

More subtle is the mistake

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = 0$$

which looks OK, but it isn't: it's still an equation that says that some vector and some number are equal. In this case the vector $\left(\begin{smallmatrix} 0 \\ 0 \\ 0 \end{smallmatrix}\right)$ is the ***zero vector*** for which we use the symbol $\vec{\mathbf{0}}$. So

$$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \vec{\mathbf{0}}$$

is correct.

Object	**Notation**
Point	Upper case letters, A, B, etc.
Position vector	Lowercase letters with an arrow on top. The position vector $\overrightarrow{OA}$ of the point A should be $\vec{a}$, so that letters match across changes from upper to lower case.
Coordinates of a point	The coordinates of the point A are the same as the components of its position vector $\vec{a}$: we use lower case letters with a subscript to indicate which coordinate we have in mind: (a_1, a_2).

Table 1. The different notations that we have used in this chapter

9. Problems–Computing and drawing vectors

1. Simplify the following

$$\vec{a} = \begin{pmatrix}1\\-2\\3\end{pmatrix} + 3\begin{pmatrix}0\\1\\3\end{pmatrix};$$

$$\vec{b} = 12\begin{pmatrix}1\\1/3\end{pmatrix} - 3\begin{pmatrix}4\\1\end{pmatrix};$$

$$\vec{c} = (1+t)\begin{pmatrix}1\\1-t\end{pmatrix} - t\begin{pmatrix}1\\-t\end{pmatrix},$$

$$\vec{d} = t\begin{pmatrix}1\\0\\0\end{pmatrix} + t^2\begin{pmatrix}0\\-1\\2\end{pmatrix} - \begin{pmatrix}0\\0\\1\end{pmatrix}.$$

2. If $\vec{a}$, $\vec{b}$, $\vec{c}$ are as in the previous problem, then which of the following expressions mean anything? Compute those expressions that are well defined.

(a) $\vec{a} + \vec{b}$ **(b)** $\vec{b} + \vec{c}$ **(c)** $\pi\vec{a}$

(d) $\vec{b}^2$ **(e)** $\vec{b}/\vec{c}$ **(f)** $\|\vec{a}\| + \|\vec{b}\|$

(g) $\|\vec{b}\|^2$ **(h)** $\vec{b}/\|\vec{c}\|$

3. Let $\vec{a} = \begin{pmatrix}1\\-2\\2\end{pmatrix}$ and $\vec{b} = \begin{pmatrix}2\\-1\\1\end{pmatrix}$.

Compute:

(a) $||\vec{a}||$

(b) $2\vec{a}$

(c) $||2\vec{a}||^2$

(d) $\vec{a} + \vec{b}$

(e) $3\vec{a} - \vec{b}$

4. Let $\vec{u}, \vec{v}, \vec{w}$ be three given vectors, and suppose

$$\vec{a} = \vec{v}+\vec{w}, \quad \vec{b} = 2\vec{u}-\vec{w}, \quad \vec{c} = \vec{u}+\vec{v}+\vec{w}.$$

(a) Simplify $\vec{p} = \vec{a} + 3\vec{b} - \vec{c}$ and $\vec{q} = \vec{c} - 2(\vec{u} + \vec{a})$.

(b) Find numbers r, s, t such that $r\vec{a}+s\vec{b}+t\vec{c} = \vec{u}$.

(c) Find numbers k, l, m such that $k\vec{a} + l\vec{b} + m\vec{c} = \vec{v}$.

5. Prove the Algebraic Properties (68a), (68b), (68c), and (68d) in section 1.5.

6. **(a)** Does there exist a number x such that

$$\begin{pmatrix}1\\2\end{pmatrix} + \begin{pmatrix}x\\x\end{pmatrix} = \begin{pmatrix}2\\1\end{pmatrix}?$$

(b) Make a drawing of all points P whose position vectors are given by

$$\vec{p} = \begin{pmatrix}1\\2\end{pmatrix} + \begin{pmatrix}x\\x\end{pmatrix}.$$

(c) Do there exist a numbers x and y such that

$$x\begin{pmatrix}1\\2\end{pmatrix} + y\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}2\\1\end{pmatrix}?$$

7. Given points $A(2,1)$ and $B(-1,4)$ compute the vector $\overrightarrow{AB}$. *Is $\overrightarrow{AB}$ a position vector?*

8. Given: points $A(2,1)$, $B(3,2)$, $C(4,4)$ and $D(5,2)$. *Is $ABCD$ a parallelogram?*

9. Given: points $A(0,2,1)$, $B(0,3,2)$, $C(4,1,4)$ and D.

(a) If $ABCD$ is a parallelogram, then what are the coordinates of the point D?

(b) If $ABDC$ is a parallelogram, then what are the coordinates of the point D?

10. You are given three points in the plane: A has coordinates $(2,3)$, B has coordinates $(-1,2)$ and C has coordinates $(4,-1)$.

(a) Compute the vectors $\overrightarrow{AB}$, $\overrightarrow{BA}$, $\overrightarrow{AC}$, $\overrightarrow{CA}$, $\overrightarrow{BC}$ and $\overrightarrow{CB}$.

(b) Find the points P, Q, R and S whose position vectors are $\overrightarrow{AB}$, $\overrightarrow{BA}$, $\overrightarrow{AC}$, and $\overrightarrow{BC}$, respectively. *Make a precise drawing in figure 12.*

11. Have a look at figure 13

(a) Draw the vectors $2\vec{v} + \frac{1}{2}\vec{w}$, $-\frac{1}{2}\vec{v} + \vec{w}$, and $\frac{3}{2}\vec{v} - \frac{1}{2}\vec{w}$

(b) Find real numbers s, t such that $s\vec{v} + t\vec{w} = \vec{a}$.

(c) Find real numbers p, q such that $p\vec{v} + q\vec{w} = \vec{b}$.

(d) Find real numbers k, l, m, n such that $\vec{v} = k\vec{a} + l\vec{b}$, and $\vec{w} = m\vec{a} + n\vec{w}$.

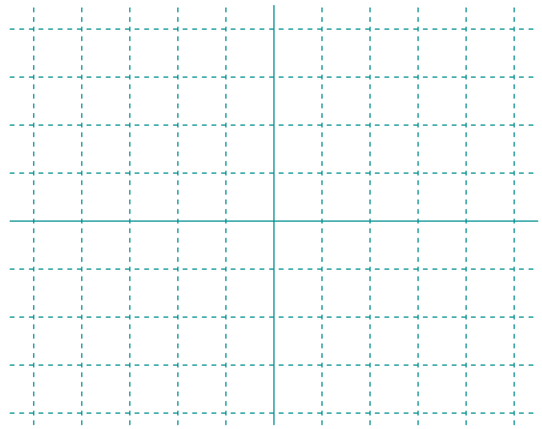
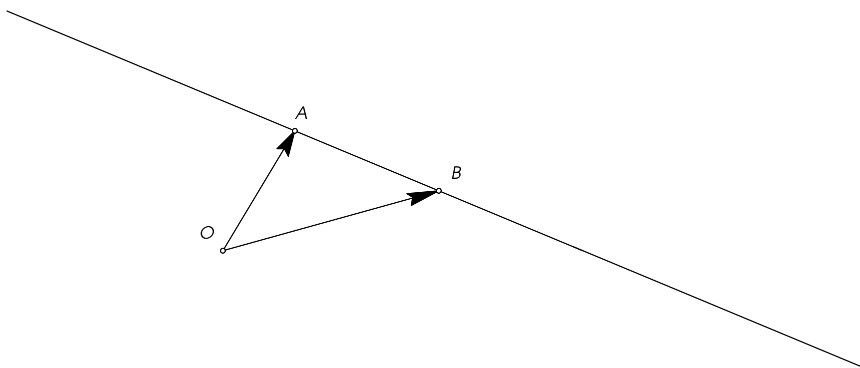


Figure 12. Your drawing for problem 9.10



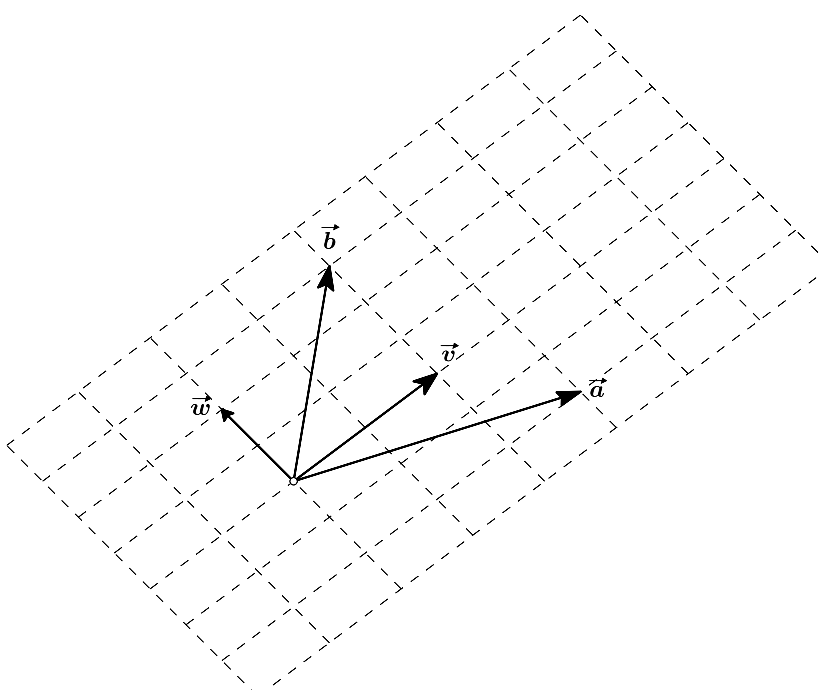


Figure 13. Drawing for problem 9.11

10. Problems–Parametric equations for a line

1. In the figure above draw the points whose position vectors are given by $\vec{\boldsymbol{x}} = \vec{\boldsymbol{a}} + t(\vec{\boldsymbol{b}} - \vec{\boldsymbol{a}})$ for $t = 0, 1, \frac{1}{3}, \frac{3}{4}, -1, 2$. (as always, $\vec{\boldsymbol{a}} = \overrightarrow{OA}$, etc.)

2. In the figure above also draw the points whose position vector are given by $\vec{\boldsymbol{x}} = \vec{\boldsymbol{b}} + s(\vec{\boldsymbol{a}} - \vec{\boldsymbol{b}})$ for $s = 0, 1, \frac{1}{3}, \frac{3}{4}, -1, 2$.

3. (a) Find a parametric equation for the line ℓ through the points $A(3,0,1)$ and $B(2,1,2)$.

(b) Where does ℓ intersect the coordinate planes?

4. (a) Find a parametric equation for the line which contains the two vectors

$$\vec{a} = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} \text{ and } \vec{b} = \begin{pmatrix} 3 \\ 2 \\ 3 \end{pmatrix}.$$

(b) The vector $\vec{c} = \begin{pmatrix} c_1 \\ 1 \\ c_3 \end{pmatrix}$ is on this line.

What is $\vec{c}$?

5. [Group Problem] Consider a triangle ABC and let $\vec{\boldsymbol{a}}, \vec{\boldsymbol{b}}, \vec{\boldsymbol{c}}$ be the position vectors of A, B, and C.

(a) Compute the position vector of the midpoint P of the line segment BC. Also compute the position vectors of the midpoints Q of AC and R of AB. (Make a drawing.)

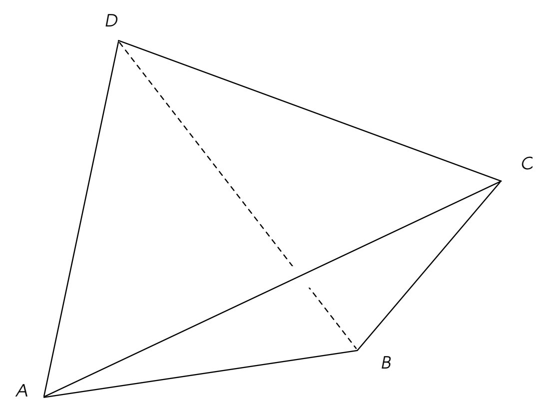
(b) Let M be the point on the line segment AP which is twice as far from A as it is from P. Find the position vector of M.

(c) Show that M also lies on the line segments BQ and CR.

6. [Group Problem] Let $ABCD$ be a tetrahedron, and let $\vec{\boldsymbol{a}}, \vec{\boldsymbol{b}}, \vec{\boldsymbol{c}}, \vec{\boldsymbol{d}}$ be the position vectors of the points A, B, C, D.

(a) Find position vectors of the midpoint P of AB, the midpoint Q of CD and the midpoint M of PQ.

(b) Find position vectors of the midpoint R of BC, the midpoint S of AD and the midpoint N of RS.



11. Problems–Orthogonal decomposition of one vector with respect to another

1. Given the vectors $\vec{\boldsymbol{a}} = \left(\begin{smallmatrix} 2 \\ 1 \\ 3 \end{smallmatrix}\right)$ and $\vec{\boldsymbol{b}} = \left(\begin{smallmatrix} 1 \\ 1 \\ 0 \end{smallmatrix}\right)$ find $\vec{\boldsymbol{a}}^{/\!/}, \vec{\boldsymbol{a}}^{\perp}, \vec{\boldsymbol{b}}^{/\!/}, \vec{\boldsymbol{b}}^{\perp}$ for which

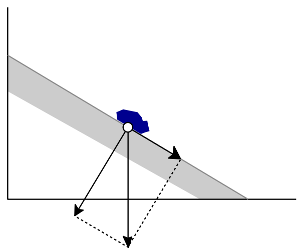
$$\vec{\boldsymbol{a}} = \vec{\boldsymbol{a}}^{/\!/} + \vec{\boldsymbol{a}}^{\perp}, \text{ with } a^{/\!/} /\!/ \vec{\boldsymbol{b}}, a^{\perp} \perp \vec{\boldsymbol{b}},$$

and

$$\vec{\boldsymbol{b}} = \vec{\boldsymbol{b}}^{/\!/} + \vec{\boldsymbol{b}}^{\perp}, \text{ with } b^{/\!/} /\!/ \vec{\boldsymbol{a}}, b^{\perp} \perp \vec{\boldsymbol{a}}.$$

2. Bruce left his backpack on a hill, which in some coordinate system happens to be the line with equation $12x_1 + 5x_2 = 130$.

The force exerted by gravity on the backpack is $\vec{\boldsymbol{f}}_{\text{grav}} = \left(\begin{smallmatrix} 0 \\ -mg \end{smallmatrix}\right)$. Decompose this force into a part perpendicular to the hill, and a part parallel to the hill.



3. An eraser is lying on the plane $\mathcal{P}$ with equation $x_1 + 3x_2 + x_3 = 6$. Gravity pulls the eraser down, and exerts a force given by

$$\vec{f}_{\text{grav}} = \begin{pmatrix} 0 \\ 0 \\ -mg \end{pmatrix}.$$

(a) Find a normal $\vec{n}$ for the plane $\mathcal{P}$.

(b) Decompose the force $\vec{f}$ into a part perpendicular to the plane $\mathcal{P}$ and a part perpendicular to $\vec{n}$.

12. Problems–The dot product

1. **(a)** Simplify $\|\vec{a} - \vec{b}\|^2$.

(b) Simplify $\|2\vec{a} - \vec{b}\|^2$.

(c) If $\vec{a}$ has length 3, $\vec{b}$ has length 7 and $\vec{a} \bullet \vec{b} = -2$, then compute $\|\vec{a} + \vec{b}\|$, $\|\vec{a} - \vec{b}\|$ and $\|2\vec{a} - \vec{b}\|$.

2. Simplify $(\vec{a} + \vec{b}) \bullet (\vec{a} - \vec{b})$.

3. Find the lengths of the sides, and the angles in the triangle ABC whose vertices are $A(2,1)$, $B(3,2)$, and $C(1,4)$.

4. **[Group Problem]** Given: $A(1,1)$, $B(3,2)$ and a point C which lies on the line with parametric equation $\vec{c} = \left(\begin{smallmatrix} 0 \\ 3 \end{smallmatrix}\right) + t\left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$. If $\triangle ABC$ is a right triangle, then where is C? (There are three possible answers, depending on whether you assume A, B or C is the right angle.)

5. **(a)** Find the defining equation and a normal vector $\vec{n}$ for the line ℓ which is the graph of $y = 1 + \frac{1}{2}x$.

(b) What is the distance from the origin to ℓ?

(c) Answer the same two questions for the line m which is the graph of $y = 2 - 3x$.

(d) What is the angle between ℓ and m?

6. Let ℓ and m be the lines with parametrizations

$$\ell : \vec{x} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} + t\begin{pmatrix} 1 \\ 2 \end{pmatrix},$$
$$m : \vec{x} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} + s\begin{pmatrix} -2 \\ 3 \end{pmatrix}$$

Where do they intersect, and find the angle between ℓ and m.

7. Let ℓ and m be the lines with parametrizations

$$\ell : \vec{x} = \begin{pmatrix} 2 \\ 1 \\ -4 \end{pmatrix} + t\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix},$$
$$m : \vec{x} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} + s\begin{pmatrix} -2 \\ 0 \\ 3 \end{pmatrix}$$

Do ℓ and m intersect? Find the angle between ℓ and m.

8. Let ℓ and m be the lines with parametrizations

$$\ell : \vec{x} = \begin{pmatrix} 2 \\ \alpha \\ 1 \end{pmatrix} + t\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix},$$
$$m : \vec{x} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} + s\begin{pmatrix} -2 \\ 0 \\ 3 \end{pmatrix}$$

Here α is some unknown number.

If it is known that the lines ℓ and m intersect, what can you say about α?

13. Problems–The cross product

1. Compute the following cross products

(a) $\begin{pmatrix}3\\1\\2\end{pmatrix} \times \begin{pmatrix}3\\2\\1\end{pmatrix}$

(b) $\begin{pmatrix}12\\-71\\3\frac{1}{2}\end{pmatrix} \times \begin{pmatrix}12\\-71\\3\frac{1}{2}\end{pmatrix}$

(c) $\begin{pmatrix}1\\0\\0\end{pmatrix} \times \begin{pmatrix}1\\1\\0\end{pmatrix}$

(d) $\begin{pmatrix}\sqrt{2}\\1\\0\end{pmatrix} \times \begin{pmatrix}0\\\sqrt{2}\\0\end{pmatrix}$

2. Compute the following cross products

(a) $\vec{\imath}\times(\vec{\imath} + \vec{\jmath})$

(b) $(\sqrt{2}\vec{\imath} + \vec{\jmath})\times\sqrt{2}\vec{\jmath}$

(c) $(2\vec{\imath} + \vec{k})\times(\vec{\jmath} - \vec{k})$

(d) $(\cos\theta\,\vec{\imath} + \sin\theta\,\vec{k})\times(\sin\theta\,\vec{\imath} - \cos\theta\,\vec{k})$

3. **(a)** Simplify $(\vec{a} + \vec{b})\times(\vec{a} + \vec{b})$.

(b) Simplify $(\vec{a} - \vec{b})\times(\vec{a} - \vec{b})$.

(c) Simplify $(\vec{a} + \vec{b})\times(\vec{a} - \vec{b})$.

4. *True or False:* If $\vec{a}\times\vec{b} = \vec{c}\times\vec{b}$ and $\vec{b} \neq \vec{0}$ then $\vec{a} = \vec{c}$?

5. **[Group Problem]** Given $A(2,0,0)$, $B(0,0,2)$ and $C(2,2,2)$. Let $\mathcal{P}$ be the plane through A, B and C.

(a) Find a normal vector for $\mathcal{P}$.

(b) Find a defining equation for $\mathcal{P}$.

(c) What is the distance from $D(0,2,0)$ to $\mathcal{P}$? What is the distance from the origin $O(0,0,0)$ to $\mathcal{P}$?

(d) Do D and O lie on the same side of $\mathcal{P}$?

(e) Find the area of the triangle ABC.

(f) Where does the plane $\mathcal{P}$ intersect the three coordinate axes?

6. **(a)** Does $D(2,1,3)$ lie on the plane $\mathcal{P}$ through the points $A(-1,0,0)$, $B(0,2,1)$ and $C(0,3,0)$?

(b) The point $E(1,1,\alpha)$ lies on $\mathcal{P}$. What is α?

7. Given points $A(1,-1,1)$, $B(2,0,1)$ and $C(1,2,0)$.

(a) Where is the point D which makes $ABCD$ into a parallelogram?

(b) What is the area of the parallelogram $ABCD$?

(c) Find a defining equation for the plane $\mathcal{P}$ containing the parallelogram $ABCD$.

(d) Where does $\mathcal{P}$ intersect the coordinate axes?

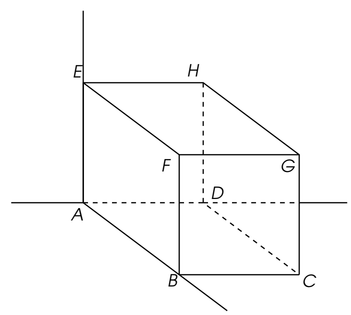
8. Given points $A(1,0,0)$, $B(0,2,0)$ and $D(-1,0,1)$ and $E(0,0,2)$.

(a) If $\mathfrak{P} = \genfrac{}{}{0pt}{}{ABCD}{EFGH}$ is a parallelepiped, then where are the points C, F, G and H?

(b) Find the area of the base $ABCD$ of $\mathfrak{P}$.

(c) Find the height of $\mathfrak{P}$.

(d) Find the volume of $\mathfrak{P}$.



9. **[Group Problem]** Let $\genfrac{}{}{0pt}{}{ABCD}{EFGH}$ be the cube with A at the origin, $B(1,0,0)$, $D(0,1,0)$ and $E(0,0,1)$.

(a) Find the coordinates of all the points A, B, C, D, E, F, G, H.

(b) Find the position vectors of the midpoints of the line segments AG, BH, CE and DF. Make a drawing of the cube with these line segments.

(c) Find the defining equation for the plane BDE. Do the same for the plane CFH. Show that these planes are parallel.

(d) Find the parametric equation for the line through AG.

(e) Where do the planes BDE and CFH intersect the line AG?

(f) Find the angle between the planes BDE and BGH.

(g) Find the angle between the planes BDE and BCH. Draw these planes.