Calculus
for the Forgetful

a short book on

how to understand more and memorize less

Calculus for the Forgetful

Wojciech Kosek

March 4, 2007

Contents

Preface

The purpose of this short book is to provide the reader with a concise treatment of single variable calculus. The main emphasis is on the understanding of ideas and facts, rather than their memorization. Informal, intuitive explanations are included for most facts, whenever feasible. Most of us tend to forget the fine details of proofs, but it is good to have a general understanding of why various statements are true. My goal is to help the reader understand the main ideas of calculus. There are certainly many topics, which are omitted for the sake of brevity. This book does not aim to be an encyclopedic treatment of the subject. Instead, I try to stay with the core ideas, with just a few digressions here and there.

The book can be used as a supplemental text by students currently taking a calculus course. It is also for those who took calculus in the past and who would like to organize their knowledge, for example while taking another class in which calculus is needed. Those who plan to take the GRE could also benefit from reading the book. The size of it is small enough to be carried around to other classes in which calculus is used. There are many exercises and over 100 examples in the text and some more challenging problems at the end of each Chapter. The problems are designed as an enhancement, but are not sufficient for most people to acquire proficiency in calculus.

I emphasize the core ideas and concepts, but include occasional digressions and comments, which show possible avenues for further exploration. Many of these side trips are in smaller print and skipping them will not cause the reader any difficulties in comprehending the rest. It is my hope that the reader will not feel intimidated by any of the nonessential material.

Everyone has a somewhat different learning style. It is my hope that this book will make calculus more enjoyable and easier to grasp for at least some readers. Chapter 1 contains a review of some basic facts and concepts used in calculus, including trigonometry and logarithms, which are a frequent source of headaches for young calculus apprentices. In Chapters 2 and 3 the concept of the derivative and methods of finding it are discussed. A limited selection of applications of derivatives is the topic of Chapter 4. Next, the concept of the integral, methods of integration and some applications are discussed in Chapters 5, 6 and 7. The book ends, in Chapter 8, with a brief introduction to infinite series.

In addition, there are four appendices, which provide some additional review of the background material (A), a summary of the book (B), answers to selected problems (C) and a few proofs, which are neat, but not essential for the general understanding of the subject (D).

I wish to thank all those who helped me make this book better, especially my students who were subjected to earlier versions of this work and who generously shared their insights with me. Special thanks go to Andrea Buchwald and Chris Kempes for their diligent proofreading and helpful comments. I am grateful to Courtney Gibbons for all her help in making the figures.

Observing all those who taught me mathematics and other subjects is how I developed my own teaching style. I owe a debt of gratitude to all of my teachers including my parents, my high school teacher, Mrs. Maria Czerska-Lazarowicz; my undergraduate advisors, Professors Roman Ger, Jerzy Klamka and Andrzej Świerniak; my Ph.D. advisor, Professor Isaac Kornfeld and countless faculty and colleagues of mine, from whom I learned mathematics and how to explain it. Most of all, I wish to thank my wife Iwona and our daughter Magdalena, whose support and infinite patience made this book possible.

Wojciech K Kosek, Colorado College

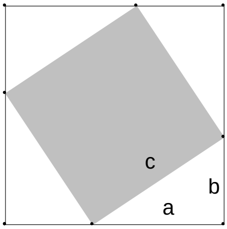
Chapter 1

Preliminary Concepts and Facts

1.1 Real Numbers

The ancient Greeks discovered that the length of the diagonal of a square of side 1 has the length of $\sqrt{2}$. This follows from what is known as the Pythagorean Theorem: the sum of the areas of the squares built on the sides of the right triangle is the same as the area of the square built on the hypotenuse. This is clear from the picture: the area of the big square can be calculated in two ways: $(a+b)^2 = c^2 + 4 \cdot \frac{1}{2}ab$. Expanding the left hand side gives:

$$a^2 + b^2 = c^2.$$



What is the big deal about it? The Greeks believed that all numbers in nature can be represented as fractions of integers. After specifying two points on a straight line as 0 and 1, it is natural to identify every point on this line with a number: we call such numbers real. However, $\sqrt{2}$ *cannot* be represented as a ratio of two integers $\frac{m}{n}$. (See Appendix D.1 for the proof). This was a serious problem: if mathematicians do not even know what a number is, what do they know? By now, we have come to terms with this realization: there are actually more real numbers which are not fractions of integers (irrationals) than there are those which can be written in that form. (We also realize that there are a lot more questions for which we do not have answers than those for which we do...)

The following is the standard classification of numbers along with notation commonly used to denote them:

$\mathbb{N}$: the natural numbers; whole positive integers and zero.

$\mathbb{Z}$: the integers; positive, negative and zero:

$\mathbb{Q}$: the rational numbers; fractions of integers.

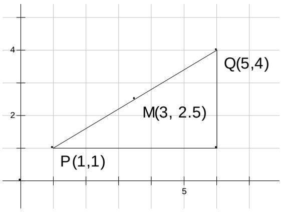
$\mathbb{R}$: all real numbers

In modern times people came to appreciate an even larger set of numbers, the complex numbers $\mathbb{C}$, but we need to stay focused on calculus and this issue is beyond our current interest.

Another common practice is the use of the interval notation: $(a, b]$ denotes the set of real numbers x, such that $a < x \leq b$; while $(-\infty, a]$ stands for the set of real numbers x such that $x \leq a$, and so on.

We use the symbol: $|x|$ to denote the absolute value of x, which should be thought of as the distance from 0 to x. In a similar fashion, the distance between numbers a and x is $|x - a|$.

1.2 Coordinates on the Plane



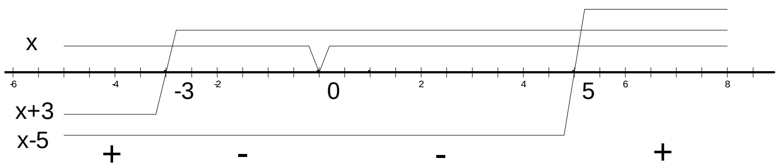
Just as there is a one-to-one correspondence between real numbers and the points on the line, we can also identify ordered pairs of real numbers with points on the plane. The distance between two points $P(x_1, y_1)$ and $Q(x_2, y_2)$ is calculated using the Pythagorean Theorem:

$$d = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2},$$

Example 2 *Solve the inequality:*

$$\frac{x^2\,(x+3)}{(x-5)^3} \geq 0. \tag{1.7}$$

First, let us notice, that the sign of either a product or a quotient of numbers does not depend on their actual values but only on their signs. The expression in inequality 1.7 has three "components": x^2, $x+3$ and $x-5$. Each component has a threshold value of x, for which it may change its sign. We need to keep track of the signs of each of the components and count how many of them are negative. Some people like to make a chart, but for me the quickest way is to draw a quick sketch. In this example, there are three threshold values: -3, 0 and 5, which in turn cut the real line into four intervals: $(-\infty,-3)$, $(-3,0)$, $(0,5)$ and $(5,\infty)$. The inequality is either true or false for all x inside each of the intervals. One way to indicate the sign of each expression is to sketch



a lines, above or below the x-axis. For example, $x+3$ gives a line below the x-axis, for $x<-3$ and above it for $x>-3$. We do the same for $(x-5)^3$, since it has the same sign as $x-5$. The line corresponding to x^2 stays above the x-axis at all times, except for $x=0$.
It is now easy to mark the sign of the whole rational expression in each of the intervals by counting the number of lines below the x-axis. Finally, we need to decide whether to include any of the threshold values in the solution. The key issue is what is their origin: the numbers which make the denominator zero can never be included, as the expression is undefined (in our case, we do not include 5). Since the inequality was not sharp ($>$), we should include 0 and -3. Therefore, the solution set to inequality 1.7 is:

$$(-\infty,-3]\cup\{0\}\cup(5,\infty)\,.$$

In general, the roots of a rational function are those values of x, for which the numerator is equal to zero and the denominator is not. That is because if the denominator is zero, the fraction is undefined.

On the other hand, the zeroes of the denominator are candidates for vertical asymptotes. However, there may or may not be an asymptote at the value for which both numerator and denominator vanish. For example, consider a function $f(x) = \frac{(x-1)(x+2)}{x-1}$, which is simply $x+2$, as long as $x \neq 1$. A few examples of graphs of rational functions are provided in Problem 4.

1.6 Exponentials and Logarithms

Exponential and logarithmic functions are of great importance in mathematics. Let us first review some facts about raising numbers to different exponents. While one can certainly raise negative numbers to integer powers, and to some fractional powers as well, we will restrict our attention to positive values of the base b. Let us pretend for a moment that we do not know how exponents work and we are attempting to come up with sensible definitions.

For a positive number b and a natural number n, we define $b^n = b \cdot b \cdot \cdots \cdot b$, as a product of "$n$ copies of b". It is easy to verify that this implies

$$b^m \cdot b^n = b^{m+n}.$$

One way to look at this is that the operation of addition among the positive integers m and n is "translated" into multiplication of b^m and b^n. This seems like a nice property and we would like to keep it when we extend the definition of raising to the power to other exponents. We would like the property $b^\alpha \cdot b^\beta = b^{\alpha+\beta}$, to remain valid for all real numbers α and β. If we could raise b to the power of 0 then we should have $b^n b^0 = b^{n+0} = b^n$, so our only choice is to define: $b^0 = 1$. This makes sense as the additive "identity" 0 is now translated to the multiplicative identity 1.

For negative exponents we want to have $b^{-n} b^n = b^{-n+n} = b^0 = 1$, thus we should define

$$b^{-n} = \frac{1}{b^n}.$$

This takes care of all of $\mathbb{Z}$, all integer exponents. How about $\mathbb{Q}$, all fractions of integers? For a fractional exponent of $\frac{1}{2}$ we should have $b^{\frac{1}{2}} b^{\frac{1}{2}} = b^{\frac{1}{2}+\frac{1}{2}} = b^1 = b$, which means that $b^{\frac{1}{2}} = \sqrt{b}$. By similar arguments we can convince ourselves that the only reasonable way to define raising b to a rational power is:

$$b^{\frac{m}{n}} = \sqrt[n]{b^m} = \left(\sqrt[n]{b}\right)^m,$$

for any integers $m, n \in \mathbb{Z}$, with $n \geq 1$.

Let us consider a function $f(x) = b^x$. So far we have established how to evaluate this function for all rational values of x. Suppose that we plot on the graph all the points (x, b^x), where $x \in \mathbb{Q}$ is rational. Then, we "connect the dots". This is possible since the set $\mathbb{Q}$ of all rationals is dense on the real number line, i.e. every real number can be approximated with any desired accuracy by the rationals. (This can be formalized by saying that the value b^α is a limit of the sequence of values $b^{\frac{m}{n}}$ for rational exponents, as the numbers $\frac{m}{n}$ approximate the irrational exponent α better and better.)

The graph of the function $f(x) = b^x$ lies above the x axis and goes through the point $(0, 1)$, for every value of b. If $b > 1$, then the function is increasing rapidly (exponentially). To obtain the graph of $f(x) = b^x$ for $b < 1$, notice that $b^x = \left(\frac{1}{b}\right)^{-x}$. All we need to do is flip the graph of $g(x) = \left(\frac{1}{b}\right)^x$ about the y-axis.

There is one value of the base b which is of particular importance: $e \approx 2.718\,3$. The symbol for the constant e honors Leonard Euler, who studied its properties extensively in the early 18th century. We will encounter this mysterious number many times in this book, but for some motivation, please feel free to look at Problems 10 and 11.

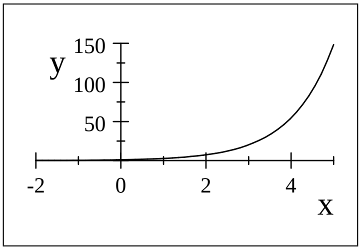


Figure 1.1: $y = e^x$

The logarithm of x to the base b is the exponent to which b needs to be raised to get x, that is

$$\log_b x = y \text{ if and only if } b^y = x.$$

In other words, $g(x) = \log_b x$ is the inverse function to the exponential function $f(x) = b^x$. For example, $10^3 = 1000$, which means that $\log_{10}(1000) = 3$. Just in case the reader forgot, the inverse function has nothing to do with a reciprocal. It is the function that brings your x back. (In mathematics we can even do that for you!)

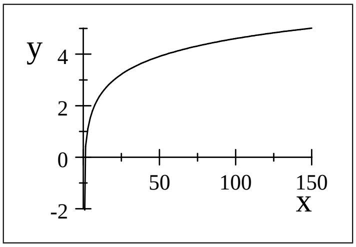


Figure 1.2: $y = \ln x$

For $b > 1$, the function $f(x) = \log_b x$ is increasing (very slowly, at least for large x). Its graph always passes through the point $(1, 0)$, regardless of the value of the base b. It is commonly accepted, at least in sciences to use the notation:

$$\begin{aligned} \log x &= \log_{10} x \quad \text{(decimal logarithm)} \\ \ln x &= \log_e x, \quad \text{(natural logarithm).} \end{aligned}$$

Logarithms have important properties, which are equivalent to the corresponding properties of the exponentials. We list them next to each other:

$$\begin{array}{ccc} \log_b(xy) = \log_b x + \log_b y & \leftrightarrows & b^\alpha b^\beta = b^{\alpha+\beta} \\ \log_b\left(\frac{x}{y}\right) = \log_b x - \log_b y & \leftrightarrows & \frac{b^\alpha}{b^\beta} = b^{\alpha-\beta} \\ \log_b(x^r) = r \log_b x & \leftrightarrows & (b^\alpha)^\beta = b^{\alpha\beta} \\ \log_b x = \frac{\log_c x}{\log_c b} = \frac{\log x}{\log b} = \frac{\ln x}{\ln b} & \leftrightarrows & (b^\alpha)^\beta = b^{\alpha\beta} \end{array} \tag{1.8}$$

where $x, y > 0$, of course. Two of the formulas correspond to the same property of exponents and in fact follow from each other. The last formula allows us to rewrite a logarithm to a different base. When working with logarithms, it is crucial to use their properties correctly. Please see Problem 9 for a list of some common mistakes.

Example 3 *Let us calculate a few values of logarithms:* $\log 10 = 1$, *since* $10^1 = 10$; $\log 100 = 2$, *since* $10^2 = 100$, $\log 1000 = 3$, *as* $10^3 = 1000$. *We can see that* $3 = \log 1000 = \log(100 \cdot 10) = \log 100 + \log 10 = 2 + 1$, *which illustrates the first property 1.8. As further examples, we can notice that* $\ln e = 1$, $\log_b 1 = 0$, *regardless of the base* b, $\log_b(b^x) = x$, $\ln(e^x) = x = e^{\ln x}$. *On the other hand* $e^{2\ln x} = \left(e^{\ln x}\right)^2 = x^2$ *(and it is not* $2x$*).*

Exercise 4 *Convince yourself that the formulas 1.8 are correct. For instance, suppose that* $\log_b x = \alpha$ *and* $\log_b y = \beta$. *Check that* $b^{\alpha+\beta} = xy$. *Rewrite it back into logarithms. For the change of base formula, start with* $y = \log_b x$, *which means* $x = b^y$. *Then take* $\log_c$ *of both sides.*

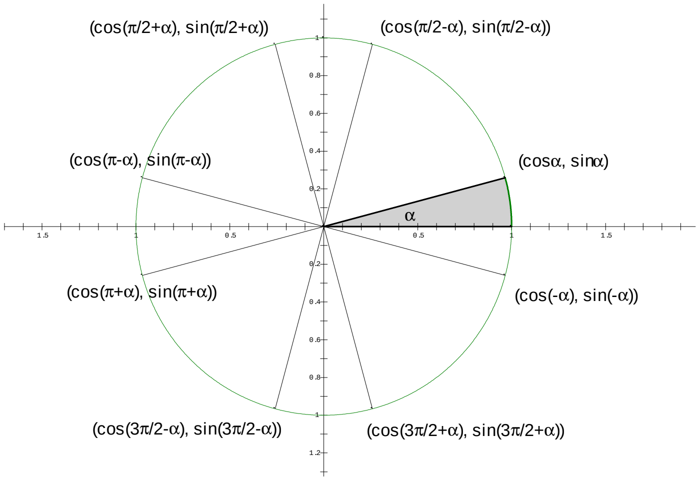


Figure 1.3: Reduction formulas for sine and cosine.

1.7 Trigonometric Functions

Trigonometry plays an important role in calculus, not just because we may be interested in some geometric considerations. Trigonometric functions are useful in integration techniques, (section 6.4). Also, the inverse trigonometric functions are antiderivatives of fairly simple functions, which we would otherwise not be able to integrate, (section 3.4).

The values of $\cos\alpha$ and $\sin\alpha$, for any real number α, are defined as the coordinates of the point on the unit circle at distance α units from $(1, 0)$ measured counterclockwise along the arc. This way of looking at the sin and cos functions explains immediately what is the sign of each function in every quadrant. It also explains several reduction identities as shown on Figure 1.3. For example, to obtain the reduction formula $\sin\left(\alpha + \frac{\pi}{2}\right) = \cos\alpha$, all we need to do is compare the coordinates of appropriate points on the unit circle. Similarly, $\sin(-\alpha) = -\sin\alpha$ and $\cos(-\alpha) = \cos\alpha$, which means that sine is an odd function while cosine is even. The reduction formulas are useful, but

Chapter 2

Limits and Derivatives, the concept.

2.1 Slope of a curve - rate of change.

One of the central ideas in calculus is that a typical function locally looks like a straight line. In other words, if we "zoom in" on the graph of a "decent" function sufficiently close, it becomes indistinguishable from a straight line. This allows us to generalize the concept of a slope of a straight line to the *slope of a curve.* Specifically, the slope of a curve $y = f(x)$ at $x = x_0$ is the slope of the line tangent to the graph of f at the point $(x_0, f(x_0))$. In other words, the **derivative** $f'(a)$ is the instantaneous rate of change of the function f at a given value of $x = a$. We are taking for granted that we know what the tangent line is. We will redefine the derivative, without using the tangent line in section 2.2.

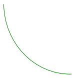
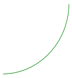
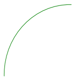
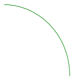
A function is called differentiable (at a point x) if it has derivative (at the point x). Such functions are locally well approximated by tangent lines. The simplest and most common reason for which a continuous function may not be differentiable is when its graphs has a "cusp". For example $f(x) = |x|$ is not differentiable at $x = 0$ (sketch a graph). A typical function in calculus is differentiable, except possibly at a few points.

What is "typical" is another story. Just like most numbers on the real line are not rational, most functions are not even continuous, not to mention differentiable. In fact, a randomly selected function is not likely to be continuous even at a single point. Moreover, among continuous functions, most are not differentiable anywhere, but this is a whole other story. Luckily, most functions which we encounter in sciences are much more smooth.

The positive (negative) value of the derivative f' means that the function itself is increasing (decreasing), at least at the immediate vicinity of the point in question.

The value of the second derivative $f'' = (f')'$, that is the derivative of the derivative, controls the rate of change of the first derivative f'. Suppose that $f''(x) > 0$, on some interval. Then the first derivative f' (or slope of f) increases, which in turn implies that the graph turns upwards. We say that the function f is **concave up**. The graph of such a function lies locally above the tangent line.

Conversely, if $f''(x) < 0$ then the function f is **concave down**, (the graph turns to the right, or downwards; the graph of f is locally below the tangent line). This is summarized below:

f'	$-$	$+$	$+$	$-$
f''	$+$	$+$	$-$	$-$
f				

(2.1)

A number x_0 in the domain of f is called:

- a **stationary** point of f if $f'(x_0) = 0$. (some authors call those critical points)
- a **critical** point of f if it is stationary or if $f'(x_0)$ does not exist
- a **local (global) maximum** point of f is $f(x) \leq f(x_0)$ in the immediate vicinity of x_0 (entire domain of f).
- a **local minimum (global)** point of f if $f(x) \geq f(x_0)$ in the immediate vicinity of x_0 (entire domain of f).
- an **inflection** point if the concavity of f changes at x_0.

It is natural in this context to consider simple harmonic motion, that is a motion of a mass on a spring with no friction. If y is a position of an object at the time t, then the Newton's Second Law of Motion implies: $y^{''} = F/m$, where F is the force acting on the mass m. On the other hand, Hooke's Law gives $F = -k_0 y$, which leads to a differential equation $y^{''} = -ky$, where $k = k_0/m$. For any constants a and b,

$$y = a \sin\left(\sqrt{k}x\right) + b \cos\left(\sqrt{k}x\right) \tag{3.7}$$

is a solution to this differential equation. By choosing the right values of the coefficients a and b, we can make sure that the initial conditions are satisfied (see problems 6 and 7 at the end of this chapter).

We end this section by stating that:

$$\begin{array}{ll} (\tan x)^{'} = \sec^2 x & (\sec x)^{'} = \sec x \tan x, \\ (\cot x)^{'} = -\csc^2 x & (\csc x)^{'} = -\csc x \cot x, \end{array} \tag{3.8}$$

all of which can be verified using the Quotient Rule from the next section.

3.2 Algebraic combinations of functions

3.2.1 Product and quotient rules

In order to see how to differentiate a product of two functions, let us consider a rectangle whose sides are some two differentiable functions of time, say $f(t)$ and $g(t)$, as in figure 3.1. The area of the rectangle $A(t) = f(t) \cdot g(t)$. When both sides of the rectangle are changing in time, the area $A(t)$ is growing "on two fronts". Let us assume that the left-lower corner of the rectangle does not move. The rate with which the right edge of the rectangle is sweeping the area is $f'(t) \cdot g(t)$. Similarily, the rate with which the area is swept by the upper edge of the rectangle is $f(t) \cdot g'(t)$. This leads to the Product Rule:

$$(fg)^{'} = f^{'}g + fg^{'}. \tag{3.9}$$

Formally speaking, let us calculate the increase of the area in a small increment of time Δt:

$$\Delta A = A(t + \Delta t) - A(t) = \Delta f(t) \cdot g(t + \Delta t) + f(t) \cdot \Delta g(t),$$

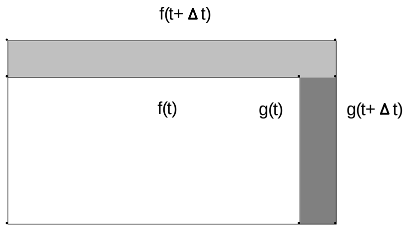


Figure 3.1: Product Rule

where $\Delta f(t) = f(t+\Delta t) - f(t)$ and $\Delta g(t) = g(t+\Delta t) - g(t)$. Dividing by Δt we get

$$\begin{aligned} \frac{\Delta A}{\Delta t} &= \frac{\Delta f(t)\cdot g(t+\Delta t) + f(t)\cdot \Delta g(t)}{\Delta t} = \\ &= \frac{\Delta f(t)}{\Delta t}\cdot g(t+\Delta t) + f(t)\cdot \frac{\Delta g(t)}{\Delta t} \end{aligned}$$

and taking the limit as $\Delta t \to \infty$ we obtain the product rule (3.9).

Having established the product rule by the above argument, let us use it to verify the quotient rule:

$$\left(\frac{f}{g}\right)' = \frac{f' g - f g'}{g^2}. \tag{3.10}$$

Suppose that a function u is a quotient $u = \frac{f}{g}$. Then $f = u\cdot g$, hence by the product rule $f' = u' g + u g'$. Solving this equation for u' we obtain

$$u' = \frac{f' - u g'}{g} = \frac{f' - \frac{f}{g} g'}{g} = \frac{f' g - f g'}{g^2}.$$

Example 53 *To find the area under the graph of* $g(x) = \frac{1}{x^2+1}$ *over* $[0,1]$ *we calculate:* $\int_0^1 \frac{1}{x^2+1} dx = \arctan x|_0^1 = \frac{\pi}{4}$.

Example 54 *Let* $F(x) = \int_0^x \frac{\sin t}{t} dt$. *The task of finding an antiderivative of* $\frac{\sin t}{t}$ *is an impossible one, at least as a so called elementary function. The reader can easily verify by differentiation, that any simple pretender for an antiderivative of* $\frac{\sin t}{t}$ *is an usurper! However, we can still find the derivative* $\frac{d}{dx} F(x) = \frac{\sin x}{x}$, *by the FTC, Version I.*

The Fundamental Theorem of Calculus makes it worth our while to find some methods for finding antiderivatives of functions. Compared to differentiation (finding derivatives) the process of integration (finding antiderivatives) is much harder. The good news is, that there are many calculators and software packages capable of symbolic integration. This makes it somewhat less important to be able to do it all on paper. However, it is still useful to acquire some degree of proficiency in integration techniques.

5.3 Common integration mistakes

Let us begin with some good news. All the formulas for derivatives can be applied "in reverse". There is little point in memorizing various formulas of that sort, as long as we know how to differentiate and can reverse the process.

Example 55 *We know that* $\frac{d}{dx}\left(x^4\right) = 4x^3$. *Therefore* $\int 4x^3 dx = x^4 + C$. *In order to integrate* x^α, *one needs to increase the exponent by one and multiply the result by the reciprocal of the new exponent:*

$$\int x^\alpha dx = \frac{1}{\alpha+1} x^{\alpha+1} + C$$

if $\alpha \neq -1$. *For* $\alpha = -1$ *we have of course:* $\int x^{-1} dx = \ln x + C$.

Clearly, integration is more delicate than differentiation. There is some good news, integration is a linear operation:

$$\begin{aligned}\int f(x)+g(x)\,dx &= \int f(x)\,dx+\int g(x)\,dx,\\ \int k\cdot f(x)\,dx &= k\cdot\int f(x)\,dx.\end{aligned}$$

Sooner or later we must come to terms with some realy bad news: there are **NO RULES for how to integrate a product or a quotient of functions!** It may have taken a while to learn how to differentiate products and quotients of functions, but at least there were rules. Not so with integration. In particular:

$$\int f(x)\cdot g(x)\,dx \neq \int f(x)\,dx\cdot\int g(x)\,dx \tag{5.6}$$

and

$$\int \frac{f(x)}{g(x)}dx \neq \frac{\int f(x)\,dx}{\int g(x)\,dx}. \tag{5.7}$$

In general, it is much better for a calculus apprentice to say "I am sorry, but I am unable to find an antiderivative of this function", than to give an obviously wrong answer! Integration can be a complicated process and even the best of us make simple mistakes once in a while. However, one is well advised to check integrals by differentiation to avoid giving bad answers, which tend to frustrate those who read such erroneous solutions. The following examples are not intented to offend the reader's intelligence, but to illustrate the "antirules" (5.6), (5.7) and a few others. A more experienced reader can skip the following examples, however in author's experience, the common mistakes below are routinely committed even by good students. It takes a while to resist "obvious" anwers.

Example 56 *Since* $(e^x)'=e^x$, *we have* $\int e^x dx=e^x+C$. *However, a common mistake is to attempt integration of functions of the form* $e^{g(x)}$ *in a naive way. However, by Chain Rule* $\frac{d}{dx}\left(e^{g(x)}\right)=e^{g(x)}\cdot g'(x)\neq e^{g(x)}$. *In the case when* $g(x)=kx$, *we can easily fix the problem:* $\int e^{kx}dx=\frac{1}{k}e^{kx}+C$, *for any constant* $k\neq 0$. *Unfortunately, this does not work when* $g(x)$ *is not a linear function.*

The lesson is:

$$\int e^{f(x)}dx \neq e^{f(x)} + C \tag{5.8}$$

Example 57 *Continuing the previous example,* $\int xe^{x^2}dx = \frac{1}{2}e^{x^2} + C$*, which is easily verified by differentiation. Apparently, having the* x *in front of* e^{x^2} *is a blessing, not a curse. A common mistake which beginners make is to think as follows: "I know how to integrate* x*. If I only knew how to integrate* e^{x^2}*, maybe I would make some progress in this integral". There are two problems with that reasoning. First, even if we knew how to integrate* e^{x^2}*, this would not directly help us integrate the product* xe^{x^2}*. A second problem is that it is not possible to write integral of* e^{x^2} *in terms of elementary functions. (There are tables of values of a closely related integral in every statistics book. They would not have been there if one could just find a nice antiderivative formula...)*

Example 58 *In the same flavor: we know that* $\int \frac{1}{x}dx = \ln x + C$ *and that* $\int \cos x dx = \sin x + C$*. This does not help in finding the antiderivative of* $\frac{\cos x}{x} = \frac{1}{x} \cdot \cos x$*. It certainly is not* $(\ln x) \cdot (\sin x)$*, as* $\frac{d}{dx}((\ln x) \cdot (\sin x)) = \frac{1}{x}\sin x + \cos x \ln x$*. This illustrates 5.6.*

Example 59 *To illustrate 5.7 let us consider* $\frac{\cos x}{x}$*. Some students attempt to integrate this quotient one piece at a time:* $\int \frac{\cos x}{x}dx = \frac{\sin x}{\frac{1}{2}x^2} + C = \frac{2\sin x}{x^2} + C$*. This is of course a bad idea, since by the Quotient Rule the derivative* $\frac{d}{dx}\left(\frac{2\sin x}{x^2}\right) = \frac{2x\cos x - 4\sin x}{x^3}$*, which is not even close to* $\frac{\cos x}{x}$*.*

Example 60 *Another common error is to misuse the fact that* $(\ln x)' = \frac{1}{x}$*. Yes, that means that* $\int \frac{1}{x}dx = \ln x + C$ *(OK, it is more general to write* $\ln|x| + C$ *instead, to allow for negative values of* x*). However,*

$$\int \frac{1}{u(x)}dx \neq \ln u(x) + C \tag{5.9}$$

Differentiating the right hand side of 5.9 we get $\frac{d}{dx}\ln u(x) = \frac{1}{u(x)} \cdot u'(x) = \frac{u'(x)}{u(x)}$*. For instance, we have:*

$$\int \frac{1}{x^2+1}dx = \arctan x + C \neq \ln\left(1 + x^2\right) + C$$

Chapter 8

Infinite Sequences and Series

We know that many functions can be well approximated by polynomials, at least on intervals of finite length. For example, we know using the MacLaurin polynomial for $f(x) = e^x$ (see section 4.4) we can say that

$$e^x \approx 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} = \sum_{k=0}^{n} \frac{x^k}{k!}.$$

It would certainly be nice to be able to replace the approximate equality "$\approx$" with the exact one "$=$", that is to write

$$e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + \cdots = \sum_{k=0}^{\infty} \frac{x^k}{k!}. \tag{8.1}$$

This requires making some sense out of adding infinitely many numbers, which is denoted by the innocent looking "$+\ldots$". This is the goal of this chapter.

Example 105 *Assuming that the formula 8.1 is true and that it is legal to differentiate such an infinite "polynomial", let us find the derivative of e^x. As* *expected, we obtain:*

$$\begin{aligned}\frac{d}{dx}\left(e^x\right) &= 0+1+\frac{2x}{2}+\frac{3x^2}{3!}+\cdots+\frac{nx^{n-1}}{n!}+\dots\\ &= 1+x+\frac{x^2}{2}+\frac{x^3}{3!}+\cdots+\frac{x^n}{n!}+\cdots=e^x.\end{aligned}$$

Example 106 *Let us replace the variable x in the formula 8.1 by x^2. We get*

$$e^{x^2}=1+x^2+\frac{x^4}{2}+\frac{x^6}{3!}+\cdots+\frac{x^{2n}}{n!}+\cdots=\sum_{k=0}^{\infty}\frac{x^{2k}}{k!}.$$

Integrating term by term, we obtain

$$\begin{aligned}\int e^{x^2}dx &= \int\left(1+x^2+\frac{x^4}{2}+\frac{x^6}{3!}+\cdots+\frac{x^{2n}}{n!}+\dots\right)dx\\ &= x+\frac{x^3}{3}+\frac{x^5}{5\cdot 2}+\frac{x^7}{7\cdot 3!}+\cdots+\frac{x^{2n+1}}{(2n+1)\,n!}+\cdots+C\\ &= \sum_{k=0}^{\infty}\frac{x^{2k+1}}{(2k+1)\,k!}+C.\end{aligned}$$

This should be a nice incentive to study a fairly delicate concept of an infinite series and its convergence. There are many functions impossible to integrate in the traditional sense: we are unable to find their neatly packaged antiderivatives. However, if under some conditions, one can differentiate and integrate such infinite sums term by term, then we can at least express certain integrals in the form of infinite sums.

8.1 Sequences and their limits

An infinite sequence is an infinite list of numbers of the general form:

$$a_1, a_2, \dots, a_k, a_{k+1}, \dots.$$

Individual entries are called **terms** of the sequence. A standard notation for a sequence with the general term a_k is $\{a_k\}_{k=1}^{\infty}$, (or simply $\{a_k\}$, for short.)

Another way of looking at this concept is that an infinite sequence is a real-valued function whose domain consists of all positive integers.

We will say that the sequence $\{a_k\}_{k=1}^{\infty}$ converges to a real number L, and write

$$\lim_{k\to\infty} a_k = L$$

if a_k approaches L to within any desired tolerance as k increases without bound.

More formally, for a finite L, we say that $\lim_{k\to\infty} a_k = L$ if for every $\varepsilon > 0$, there exists an $N > 0$, such that $|a_k - L| < \varepsilon$ whenever $k > N$.

8.2 Infinite Series

Commercial: Making sense of adding infinitely many numbers would allow for example to express a function as an infinite Taylor Series, instead of just approximating it by a Taylor Polynomial!

The goal of this section is to make sense out of expressions involving adding infinitely many numbers, like:

$$\sum_{k=1}^{\infty} a_k = a_1 + a_2 + a_3 + \cdots + a_k + a_{k+1} + \ldots.$$

In other words - what could these mysterious dots at the end of the sum mean? We will use the word **series** or **infinite series** to refer to an expression of this form.

Given a sequence $\{a_k\}_{k=1}^{\infty}$ consider another sequence $\{S_n\}_{n=1}^{\infty}$, defined as a **sequence of partial sums** of the original sequence:

$$S_n = \sum_{k=1}^{n} a_k = a_1 + a_2 + a_3 + \cdots + a_n.$$

It is sensible to say that the larger value of n we take, the closer the n-th partial sum S_n is to the sum of all the terms. Formally, if the sequence S_n converges to S, we will say that the series $\sum_{k=1}^{\infty} a_k$ **converges** and will call number S the **sum** of the series $\sum_{k=1}^{\infty} a_k$, that is

$$\sum_{k=1}^{\infty} a_k = \lim_{n\to\infty} S_n = S.$$

Otherwise we say that the series **diverges**.

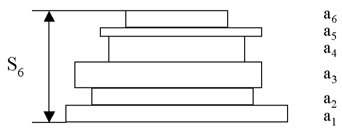
Example 107 *Consider a series* $\sum_{k=1}^{\infty} \frac{1}{2^k} = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \dots$ *Does it converge? In this case we have* $a_k = \frac{1}{2^k}$ *for all integer values of* k. *The partial sums are*

$$\begin{aligned} S_1 &= \frac{1}{2} = 1 - \frac{1}{2} \\ S_2 &= \frac{1}{2} + \frac{1}{4} = \frac{3}{4} = 1 - \frac{1}{4} \\ S_3 &= \frac{1}{2} + \frac{1}{4} + \frac{1}{8} = 1 - \frac{1}{8} \\ &\vdots \\ S_n &= \sum_{k=1}^{n} \frac{1}{2^k} = 1 - \frac{1}{2^n}, \end{aligned}$$

therefore

$$S = \lim_{n\to\infty} S_n = \lim_{n\to\infty} \left(1 - \frac{1}{2^n}\right) = 1,$$

and we will say that the series $\sum_{k=1}^{\infty} \frac{1}{2^k}$ *so the series converges to* 1. *In short, we usually just write* $\sum_{k=1}^{\infty} \frac{1}{2^k} = 1$.



Let us assume for the moment that all the terms a_k are positive. A good way to visualize the concept of the series is to imagine a stack made of infinitely many bricks of height a_k. Clearly, if there is any hope that this stack has finite height, the terms a_k must be getting smaller and smaller. That is, in order for the infinite sum $\sum_{k=1}^{\infty} a_k = \lim S_n$ to converge to *something,* it is necessary that $\lim_{k\to\infty} a_k = 0$. One other way to see that is to realize that $a_k = S_k - S_{k-1}$, for all k (or at least for $k \geq 2$). If the $\lim_{k\to\infty} S_k = S$ exists (a finite limit that is), then for sufficiently large of k, the values of S_k are close to S. Therefore, the difference $S_k - S_{k+1}$ must be small.

Once again, we have established that the limit of the general term $\lim_{k\to\infty} a_k$ must be 0 in order for the series $\sum_{k=1}^{\infty} a_k$ to converge. It is a common misconception, that this necessary condition for the convergence of a series is also sufficient. The following important example shows that *this is not case.*

Example 108 *The* ***harmonic series*** $\sum_{k=1}^{\infty} \frac{1}{k} = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \dots$ *diverges. Let us first notice that*

$$\begin{aligned}
\frac{1}{3} + \frac{1}{4} &> \frac{1}{4} + \frac{1}{4} = \frac{1}{2} \\
\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} &> \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{1}{2} \\
\frac{1}{9} + \frac{1}{10} + \frac{1}{11} + \frac{1}{12} + \frac{1}{13} + \frac{1}{14} + \frac{1}{15} + \frac{1}{16} &> 8 \cdot \frac{1}{16} = \frac{1}{2} \\
&\vdots \\
\frac{1}{2^i + 1} + \frac{1}{2^i + 2} + \frac{1}{2^i + 3} + \dots \frac{1}{2^{i+1}} &> 2^i \cdot \frac{1}{2^{i+1}} = \frac{1}{2}
\end{aligned}$$

Hence, by taking sufficiently many terms of the harmonic series we have

$$\begin{aligned}
S_{2^{i+1}} = 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \dots + \\
+ \left(\frac{1}{2^i + 1} + \frac{1}{2^i + 2} + \frac{1}{2^i + 3} + \dots \frac{1}{2^{i+1} - 1} + \frac{1}{2^{i+1}}\right) > \\
> 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \dots + \frac{1}{2} = 1 + \frac{i+1}{2},
\end{aligned}$$

(where $\frac{1}{2}$ in the sum above is repeated $i + 1$ times). For instance, adding the first $2^{10} = 1024$ terms of the series we have $S_{1024} > 1 + \frac{10}{2} = 6$. In the same way: $S_{2^{20}} = S_{1,048,576} > 1 + \frac{20}{2} = 11$, and so on. Therefore

$$\sum_{k=1}^{\infty} \frac{1}{k} = \lim_{n\to\infty} S_n = \infty,$$

which means that the harmonic series diverges to ∞.

Appendix B

Summary

B.1 Main concepts and ideas in the book

1. Limit of a function at a point. The concept of continuity and differentiability.

2. The concept of the derivative as rate of change and as the slope of the tangent line. Approximating a function by the tangent line at a point and by polynomials.

3. Rules for finding derivatives, including: power rule, derivatives of common functions, chain rule, product and quotient rules.

4. Some applications of derivatives, including l'Hôpital's Rule, optimization and related rates.

5. The Fundamental Theorem of Calculus.

6. Methods of integration.

7. Some applications of integrals, including arc length, areas between curves, volumes of solids and work.

8. Infinite series and some conditions for their convergence.

9. Representation of functions as an infinite series of simpler functions, such as polynomials or trigonometric polynomials.

B.2 What everyone should remember from calculus

1. The concept of a limit is important for many reasons, including properly defining continuity of functions and defining derivative of a function. Plugging in the value of x does not always work! If we get for example $\frac{0}{0}$, that does not mean that the limit does not exist, but that there is more work to be done.

2. The derivatives of basic functions such as power functions, exponential and logarithmic functions, as well as of at least some trigonometric and inverse trigonometric functions (sin, cos, tan, arcsin, arccos and arctan).

3. Implicit differentiation: think of both sides of the equation as functions of x. Then differentiate both sides using the known rules. Next, solve the resulting equation for y'.

4. Some limits (of the form $\frac{0}{0}$ and $\frac{\infty}{\infty}$) can be calculated using l'Hôpital's Rule: $\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f^{'}(x)}{g^{'}(x)}$. There is no rule like that for the product, which needs to rewritten as a quotient before attempting l'Hôpital's Rule: $f \cdot g = \frac{f}{1/g}$. Also, sometimes it is helpful to find the limit of the logarithm of the function first: $f(x) \to L$, whenever $\ln(f(x)) \to \ln L$.

5. In optimization problems the most difficult part is frequently the set up. It is important not to mix up the objective function with the constraint equation.

6. Related Rates: do not plug in numbers too soon. Keep the variables as variables, until after differentiating. Put the numbers in at the end.

7. Differentiable functions can be locally approximated by a tangent line - the first order approximation. Linear approximations of a function at various points are used for example in numerical methods for solving

differential equations (Euler's method). For a better approximation of a function we can use Taylor Polynomials.

8. The Intermediate Value, the Extreme Value and the Mean Value Theorems are examples of existence theorems. It is good to know what they say.

9. Definite integral is the signed area below the graph of the function over a given interval. It is formally defined as a limit of Riemann sums. The value of the definite integral can be approximated numerically using rectangles or trapezoids instead of the actual function.

10. The derivative of the area function is the integrand:

$$\frac{d}{dx}\int_a^x f(t)\,dt = f(x).$$

In other words:

$$\int_a^b f(x)\,dx = F(b) - F(a),$$

where F is any antiderivative of f.

11. Finding antiderivatives is a lot harder than finding derivatives. It may have taken some effort to learn how to differentiate products and quotients, but at least there were rules. There are no rules like that for integration! Knowing that there are such elaborate techniques for integrating products, one should never attempt to integrate products or quotients of functions "one piece at a time".

12. The Chain Rule for differentiation corresponds to the method of substitution for integration. It does not work well for every integral and the substituting variable (say u) must be chosen carefully. Make sure that all the remains of the old variable are gone before attempting to integrate with respect to the new variable. Also, do not forget that dx does not automatically become du, but that $du = u'\,dx = \frac{du}{dx}dx$.