Mathematical and Computational Methods

Physicists use math all of the time in nearly everything that they work on. Hence, it is critical that you become efficient in being able to use more advanced math to enable you to work on more advanced physics courses. The goal of this class is to transform you from a math technician to a math practitioner. Mathematicians take this one step further and actually create new math. We will not focus on how to do that at all in this class.

NOTE: The course will be available through June 14, 2021. No certificates will be issued for work completed after June 14, 2021.

James Freericks
Professor and McDevitt Chair
Georgetown University

James Freericks is a Professor of Physics who works in theoretical and computational physics focused on quantum mechanical phenomena. He has published over 200 peer-reviewed articles, an award-winning textbook, and has taught quantum mechanics to nonscientists since 1995. He is a Fellow of the American Physical Society and the American Association for the Advancement of Science. He also is the Divisional Councilor of the American Physical Society's Division of Computational Physics.

In physics, we use advanced math to describe many, many different things. We need to become practitioners of math. This is more than just knowing the technical details for how to perform mathematical manipulations. It is about properly understanding the details for the how and the why of the math. It is understanding the "big picture" for how mathematics is used to describe the physical world.

I describe this as making the transition from a technician to a practitioner. A technician is someone who learns a particular skill without thinking carefully about when it applies, how one applies it in a novel context, and what its limitations might be. Practitioners, know all that and more. They are experts in knowing how to use math as a tool to help elaborate the properties of the physical world. All physicists are practitioners of math. This course is the first step in your journey towards this goal.

We do not cover all the math you will need in your career. Just a small fraction. But by learning how to learn about math in this fashion, where we connect the math ideas to physical principles, you will be able to learn whatever more math you need on your own, or in the context of an upper-level physics class. Because the level of coverage of math is different between being a practitioner and creating new knowledge as a mathematician, we can and do cover a lot of ground in this class. This is simply because physicists need to know a lot.

The specific topics we cover is single-variable calculus, multivariable calculus, vector calculus theorems, complex numbers and complex variables (including the residue theorem), linear algebra, first-order linear and nonlinear differential equations, and second-order differential equations with constant coefficients.

WHAT DOES THE COURSE CONTENT INCLUDE?

The course is broken into a number of different themes and then within each theme, we have a number of specific topics that we cover.

In a broad brush, we provide a calculus review followed by the integral integral theorems of multivariable calculus. Then we have a short unit on complex numbers, which also develops the calculus of residues. Linear algebra follows, then differential equations and finally Fourier series.

Course Section Outline

Calculus Review

1. The Greeks and the development of the limit
2. The logarithm and the revolution of numerical computation from tables
3. The concept of the derivative and antiderivative
4. How to integrate
5. Multivariable integrals
6. Frullani and Gaussian integrals
7. Feynman integration

Vector Calculus

1. Vector fields
2. The divergence theorem

3. Stokes theorem
4. Gradient
5. Laplace's equation

Complex Analysis

1. Complex numbers
2. Analytic functions
3. Cauchy's theorem
4. Calculus of residues

Linear Algebra

1. Matrix multiplication
2. The determinant
3. Matrix inverses
4. Eigenvalues and eigenvectors
5. Abstract vector spaces with inner products

Differential Equations

1. First-order linear differential equations
2. First-order nonlinear differential equations
3. Differential equations with constant coefficients
4. Variation of parameters
5. Method of undetermined coefficients
6. Frenet-Serret apparatus

Fourier Series

1. Dirichlet theorem and Fourier summation formula
2. Poisson's theorem

WHAT WILL I LEARN IN THE COURSE?

The underlying colloquial goal is to transform from a technician to a practitioner in math. But this is vague from a learning design standpoint. So the precise learning goals are as follows:

Specific learning goals for the course

- Be able to apply techniques of calculus (learned in the first three semesters of a calculus sequence) to solve problems that arise in physics.
- Derive and use the geometric series in calculations.
- Manipulate power series expressions and employ them in physics contexts.

- Calculate Taylor polynomials/series of common functions and use them in approximating functions
- Follow the development for how one integrates polynomials, rational functions of polynomials, square roots of quadratics, rational functions with square roots of quadratics, and why the procedure cannot solve integrals with square roots of quartics.
- Solve integrals via parametric methods (differentiating under the integral sign) including techniques for introducing the parameter into the integrand
- Set up and integrate multidimensional integrals with variable mass density and for moments of inertia.
- Solve problems in multivariable integrals via the different integral theorems
- Solve Laplace's equation for simple geometries
- Compute the curl or divergence of a vector field and the gradient of a scalar function
- Solve complex numbers arithmetic problems
- Use calculus of residues to solve integrals
- Use row reduction to solve simultaneous linear equations
- Calculate the determinant, inverse and eigenvalues of matrices
- Construct abstract vector spaces using the definitions of a vector space and generalized inner products
- Solve any first-order linear differential equation
- Solve the six classes of nonlinear first-order differential equations
- Solve higher-order differential equations with constant coefficients (homogeneous and inhomogeneous)
- Construct representations of curves in three dimensions using the Frenet-Serret apparatus

The instructor will regularly participate in discussion boards to provide content clarification, guidance, and support.

You can also email us with important content-related questions at gux@georgetown.edu.

Mathematical and Computational Physics Lecture Notes

Jim Freericks
Professor of Physics
Georgetown University

January 2020

Devious Math Tricks

0.1 Introduction

Before jumping into the main text, we start with a primer on mathematical tools that you can use to make your mathematical work easier, faster, and more accurate. I hope you enjoy these ideas and have a chance to implement them through the course. They take some time to master, but are well worth it.

Remember that kid from algebra class who sat across from you and seemed to know the answers to every question faster and more accurately than you? I am going to let you in on a secret—most likely that "curve buster" was not a genius or born to do math—instead he or she simply learned a number of devious tricks to get through the material faster and with fewer errors. After finishing this chapter, you will be able to do so too!

The biggest myth to dispel is that geniuses in algebra have photographic memories and can recall the relevant formulas and apply them at will. Such people may exist but the brainiac across the row from you was not one of these. The reality is you need to only know something like ten important equations, and by using them as I am about to describe to you, you will find that you too will just know them when you need them. You too can become an algebra master.

0.2 Devious trick number 1 (Add zero)

$x+0=x$. This is called the "add zero" trick. It comes up so many times you cannot get through most problems without encountering it. The problem is no one has told you about it, so you often stumble onto it and fail to recognize its importance. We will see and use the add zero trick again and again. Right now, let me illustrate it for you. Consider the following problem: factor $x^2 - 25$.

Go ahead and give it a try, but do not just write down the answer if you have memorized it, see if you can derive it. Here is how the add zero trick works. An educated guess says we should add a number times x as one of the terms, the 25, suggests. We add 5x- 5x to get:

$$x^2 + 5x - 5x - 25 = (x^2 + 5x) - (5x + 25) \tag{1}$$

$$= x(x+5) - 5(x+5) \tag{2}$$

$$= (x-5)(x+5). \tag{3}$$

All we did was add zero! The creativity was to recognize what zero to add. Indeed this is how you apply the technique. You learn pretty quickly what ideas might work for what to add by practicing the method a number of times.

0.3 Devious trick number 2 (Multiply by 1)

$x \times 1 = x$. Just like the "add zero" trick, the "multiply by one" trick can be used to simplify many different expressions. It often is seen in cases where we want to rationalize a denominator, such as the following:
Rationalize the denominator of $\frac{1-\sqrt{3}}{1+2\sqrt{3}}$.

The strategy is to multiply the numerator and denominator by the same factor such that the denominator no longer has a square root. Recalling that $(a+b)(a-b) = a^2 - b^2$, our strategy is to use $1 = \frac{1-2\sqrt{3}}{1-2\sqrt{3}}$ as the term we multiply by. Now we compute:

$$\frac{1-\sqrt{3}}{1+2\sqrt{3}} \times \frac{1-2\sqrt{3}}{1-2\sqrt{3}} = \frac{1-\sqrt{3}-2\sqrt{3}+6}{1-12} = -\frac{7-3\sqrt{3}}{11}. \tag{4}$$

This "multiply by one" trick can also be used to simplify products of integers via factorials. For example:

$$(n+1)(n+2)...(2n-1)(2n) = \frac{1 \times 2 \times 3...n}{1 \times 2 \times 3...n}(n+1)(n+2)...(2n-1)(2n) = \frac{(2n)!}{n!} \tag{5}$$

0.4 Devious trick number 3 (Manipulate identities)

Remember simple identities and use them to make complicated identities. As I mentioned before, there are only a handful of identities worth remembering. Nearly everything else can be re-derived from these few. As an example, think about your trigonometry identities. There are loads upon loads of them. But they all follow from two main results—remembering the definitions of the trig functions in terms of sin and cos and using the fact that $\sin^2 + \cos^2 = 1$. Let's illustrate how. Suppose I want to express $\sec^2(\theta)$ in terms of $\tan(\theta)$. I simply recall that $\sec(\theta) = \frac{1}{\cos(\theta)}$. Then I use our "multiply by one" trick using $1 = \cos^2(\theta) + \sin^2(\theta)$, so

$$\begin{aligned} \sec^2(\theta) &= \frac{1}{\cos^2(\theta)} = \frac{1}{\cos^2(\theta)}\left(\cos^2(\theta) + \sin^2(\theta)\right) = \frac{\cos^2(\theta)}{\cos^2(\theta)} + \frac{\sin^2(\theta)}{\cos^2(\theta)} \\ &= 1 + \tan^2(\theta). \end{aligned} \tag{6}$$

0.5 Devious trick number 4 (Recognize abstracted identities)

Recognize simple identities when expressed in a more complex form. This one takes a bit more practice, as it requires you to use abstraction to recognize the simple identity. Here is an example. We already saw that $x^2 - 25 = (x+5)(x-5)$. We need to recognize it in other contexts. So

$$\frac{\sin^2(\theta) - 25}{\cos^2(\theta)} = \tan^2(\theta) - 25\sec^2(\theta) = (\tan(\theta) + 5\sec(\theta))(\tan(\theta) - 5\sec(\theta)) \tag{7}$$

will also hold. A more complicated example is to factor $(x^4 + 2x^2 - 24)$. We use the add zero trick first:

$$\begin{aligned} \left(x^4 + 2x^2 + (1-1) - 24\right) &= \left((x^2+1)^2 - 25\right) = (x^2+1+5)(x^2+1-5) \\ &= (x^2+6)(x^2-4). \end{aligned} \tag{8}$$

Note how we had to recognize $x^4 + 2x^2 + 1 = (x^2+1)^2$ was the square in the expression.

0.6 Devious trick number 5 (subtract identities)

Use information you know and change expressions to those you and know and their difference. This one is a little more difficult to recognize but is quite important. We will illustrate by example. Consider evaluating the following:

$$x^2 + x^4 + x^6 + ... = \sum_{n=1}^{\infty} x^{2n}. \tag{9}$$

This looks like a geometric series, but for x^2, not x. In other words,

$$\sum_{n=1}^{\infty} x^{2n} = \sum_{n=1}^{\infty} (x^2)^n. \tag{10}$$

This would be the geometric series except it is missing the first term. Hence,

$$\sum_{n=1}^{\infty} x^{2n} = \sum_{n=0}^{\infty} (x^2)^n - 1 = \frac{1}{1-x^2} - 1 = \frac{x^2}{1-x^2}. \tag{11}$$

We also could find it another way, by recognizing that

$$x^2 + x^4 + x^6 + ... = x^2(1 + x^2 + x^4 + ...) = x^2 \sum_{n=0}^{\infty} (x^2)^n = x^2 \frac{1}{1-x^2} = \frac{x^2}{1-x^2}, \tag{12}$$

as before.

0.7 Devious trick number 6 (index gymnastics)

We end by discussing one other skill, which involves shifting indices in summations. Students often struggle with this skill. Yet it is fairly easy to master. Let's look at the last example:

$$x^2 + x^4 + x^6 + ... = \sum_{n=1}^{\infty} x^{2n}. \tag{13}$$

We know the geometric series starts with $n = 0$, not $n = 1$. So let's shift $n \to n' + 1$. Then n' runs from 0 to ∞. We obtain:

$$\sum_{n=1}^{\infty} x^{2n} = \sum_{n'=0}^{\infty} x^{2(n'+1)} = \sum_{n'=0}^{\infty} x^{2n'} x^2 = \frac{1}{1-x^2} x^2 = \frac{x^2}{1-x^2}, \tag{14}$$

just like before. We will see that this index shifting skill becomes critically important in mastering many of the identities we develop later. I will help you recognize and develop this skill as we move forward.

0.8 Important identities

We end this section with the important identities you should know already (some others will be developed in this chapter). The main ones are the following:

1) $ax^2 + bx + c$ has two roots given by $r_\pm = -\frac{b}{2a} \pm \frac{1}{2a}\sqrt{b^2 - 4ac}$

2) $\cos^2(\theta) + \sin^2(\theta) = 1$

3) $\tan(\theta) = \frac{\sin(\theta)}{\cos(\theta)}$, $\cot(\theta) = \frac{\cos(\theta)}{\sin(\theta)}$, $\sec(\theta) = \frac{1}{\cos(\theta)}$, and $\csc(\theta) = \frac{1}{\sin(\theta)}$

4) $\cos(2\theta) = \cos^2(\theta) - \sin^2(\theta)$, $\sin(2\theta) = 2\cos(\theta)\sin(\theta)$

5) $x^2 - a^2 = (x+a)(x-a)$

6)$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$, $\sum_{n=0}^{N} x^n = \frac{1-x^{N+1}}{1-x}$

7) $\sum_{n=0}^{N} \binom{N}{n} x^{N-n} y^n = \sum_{n=0}^{N} \frac{N!}{n!(N-n)!} x^{N-n} y^n = (x+y)^N$ (the binomial theorem)

8) $e^x = \sum_{n=0}^{\infty} \frac{1}{n!} x^n$

Ok, now you are armed with the skills to tackle some tough algebra problems. Let's give it a try.

0.9 Problems

These problems are optional. But I encourage you to try them to hone your devious algebra trickery...

1. Factor $(x^2 + 2xy + y^2 - 9)$

2. Use the double angle formula, in the form $\cos(\theta) = \cos^2(\frac{\theta}{2}) - \sin^2(\frac{\theta}{2})$ to find the relation

$$\cos\left(\frac{\theta}{2}\right) = \pm\sqrt{\frac{\cos(\theta) + 1}{2}}$$

Hint: $\cos^2(\frac{\theta}{2})+\sin^2(\frac{\theta}{2}) = 1$ will help. How do you pick the correct sign?

3. Use the same double angle formula to find $\sin(\frac{\theta}{2})$. Verify your results for 2) and 3) satisfy $\cos^2(\frac{\theta}{2}) + \sin^2(\frac{\theta}{2}) = 1$

4. Show that $\sum_{n=0}^{\infty} \frac{1}{(n+1)!} x^n = \frac{e^x-1}{x}$.
   Hint: Use formula (8) above.

5. This is a challenging one. Show that

$$\sqrt{3+2\sqrt{2}} = \pm(1+\sqrt{2})$$

   by using the add zero trick to establish that the argument of the square root on the left hand side is a perfect square.

6. Use the double angle formulas for sine and cosine to show

$$\tan(2\theta) = \frac{2\tan(\theta)}{1-\tan^2(\theta)}$$

7. Use multiplication by $(1-x)$ (trick used in identity (6) above for the appropriate x) to show that:

$$1 + \sqrt[4]{2} + \sqrt{2} + \sqrt[4]{8} = \frac{1}{\sqrt[4]{2}-1}$$

8. Show that

$$\frac{(2n)!}{2^n n!} = 1 \times 3 \times 5 \times 7 \cdots (2n-3) \times (2n-1) = (2n-1)!!$$

   where the last equality defines the double factorial !! (the double factorial skips integers, so it is a product of all even or all odd integers only).

9. Show that

$$\sum_{n=1}^{\infty} (\cos(\theta))^{2n} = \frac{1}{\tan^2(\theta)}$$

   Hint: Use the geometric series.

Chapter 1

Irrational Numbers and Ratios

1.1 Introduction and Course Goals

The course will cover the following:

- Review of calculus, focusing on key concepts and ideas likely missed the first time through the material
- Multivariable calculus, Div, Grad, Curl, and multivariable integral theorems
- Complex variables and the calculus of residues
- Linear algebra, solving linear equations, eigenvectors and eigenvalues
- First and second order differential equations
- Fourier Series

Throughout the course we stress ideas in addition to learning the mechanics of how to do something. We also apply this to problems with a physics background.

1.2 Irrational Numbers

We begin with the proof of the existence of an irrational number. An irrational number is one that cannot be written as the ratio of two integers—$\frac{p}{q}$ with p and q relatively prime (have no common factors, or in "lowest-terms").

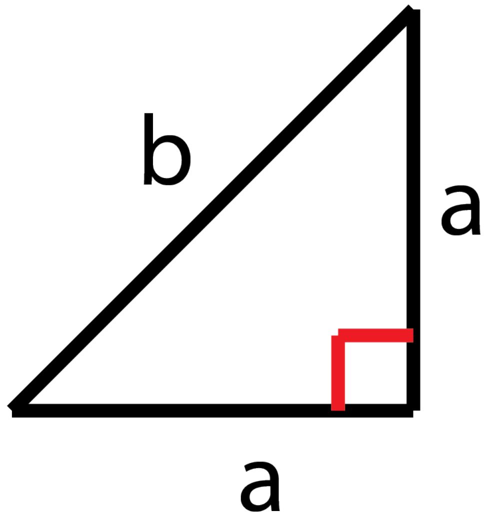

Figure 1.1: Right triangle employed in the proof. The length of each edge is a and the hypotenuse is $b = \sqrt{2}a$.

The argument is as follows: Consider the hypotenuse of an isosceles right triangle. Pythagoras says $a^2 + a^2 = b^2$ or $b^2 = 2a^2$.

Suppose we have an even integer $2n$. When we square it, it becomes $4n^2$, which is even and divisible by 4. Squaring the odd integer $2n+1$ gives $(2n+1)^2 = 4n^2 + 4n + 1$ which is an odd integer.

Now suppose $p \times e = b$ and $q \times e = a$ so that $\frac{a}{b} = \frac{q}{p}$ with p and q relatively prime and e the common factor in both a and b. Then $p^2 = 2q^2$, so p^2 is even. This means it must be the square of an even integer, so $p = 2n$. But then $2q^2$ is a multiple of 4, so $q = 2m$ for some m. Then p and q have a common factor of 2, which is not allowed, because we have set up the problem such that p and q have no common factors. Therefore the ratio $\frac{a}{b}$ cannot be written as a rational number ... it is irrational!

Irrational numbers can have odd properties. For example, if we take an irrational number and raise it to an irrational power, we can get a rational number. Here is a proof:

Consider $x = \sqrt{3}^{\sqrt{2}}$, which is an irrational raised to an irrational power. Now consider

$$x^{\sqrt{2}} = \left(\sqrt{3}^{\sqrt{2}}\right)^{\sqrt{2}} = \left(\sqrt{3}\right)^{\sqrt{2}\times\sqrt{2}} = \left(\sqrt{3}\right)^{2} = 3 \tag{1.1}$$

which is rational. Therefore either x is rational or $x^{\sqrt{2}}$ is rational. So an irrational number raised to an irrational power can yield a rational number! Note that this argument does not tell us whether the irrational is $\sqrt{3}$ or $x = \left(\sqrt{3}\right)^{\sqrt{2}}$, which when raised to an irrational power produces a rational number. (There are techniques that do answer this question, but we will not discuss further here.)

1.3 Zeno's Paradox

The book by Toeplitz discusses Zeno's Paradox — that if every step I take is half as large as the previous one, then I can never get from point A to point B because it would take an infinite number of steps.

The resolution is simple. I do need an infinite number of steps, but the total sum of all of these steps is *finite*, which is why we can arrive at point B. Next lecture, we will actually prove that the sum of all of those steps is just twice the length of the first one. This is because

$$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots = 2. \tag{1.2}$$

1.4 Ratios

What is the definition of π? Think about this before I reveal the answer.

Be sure to give it a try.

Really. Think about it.

Don't peek yet

It is the ratio of the circumference of a circle to its diameter or the ratio of the area of a circle to the square of its radius. The idea of defining π as a *ratio* comes from the Greeks. Many math concepts have hidden definitions in terms of ratios. Examples include area, sine, cosine, tangent, hyperbolic tangent, and so on.

Next we will show amazing things you can do with ratios. We will assume a well known (and in many respects self-evident) fact that the ratio of the areas of any two pie slices of the same angle of two circles is proportional to the ratio of the squares of the radii of the two circles. Consider the circle with radius 1:

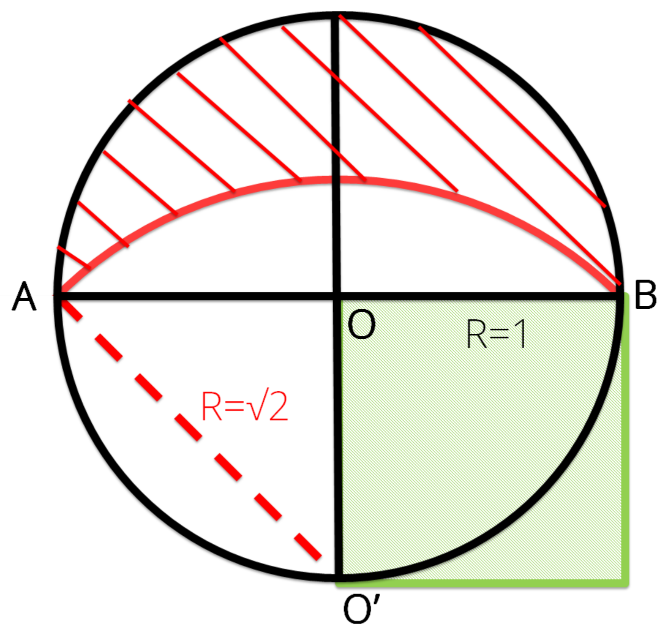

Figure 1.2: This figure has a circle of radius 1 (black) and the chord of a circle of radius $\sqrt{2}$ (red) drawn on it. The proof we will make is that the area between the two circles (hatched red) is equal to the area of the square (green).

We will prove that the red area and the green area are the same. This is close to the problem known as "squaring the circle," which has been proven to be impossible 2000 years after the Greeks gave up on their attempts. But this one can be proven, and perhaps inspired them on the bigger goal.

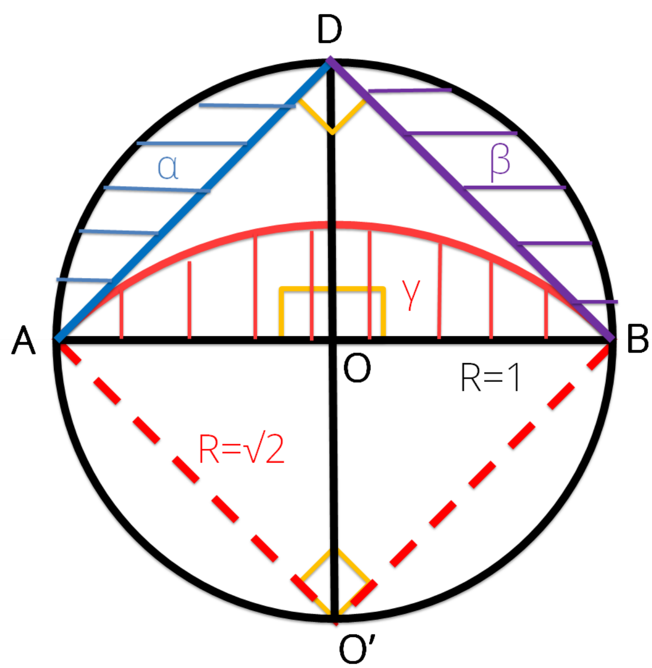

Figure 1.3: We start at the midpoint D of the upper half of the circle. Draw lines from D to A and B. This defines the hatched areas α (blue) and β (purple). We also have the hatched red area γ.

Draw straight lines from midpoint D to A and B (blue and purple). The angles AOD, DOB, and $AO'B$ are $90°$, so the the assumption we stated above implies that the ratio

$$\frac{\text{Area}\,\alpha}{\text{Area}\,\gamma} = \left(\frac{\text{rad}\,\alpha}{\text{rad}\,\gamma}\right)^2 = \frac{1}{2}. \tag{1.3}$$

A similar argument can be done for β, so we can immediately conclude that the areas satisfy $\alpha+\beta = \gamma$. Now consider $\triangle ADB$. Its area is $\frac{1}{2}\times\sqrt{2}\times\sqrt{2} = 1$. So $1 + \alpha + \beta = \frac{\pi}{2}$, which is the area of the semicircle with radius 1. Hence,

$$\frac{\pi}{2} - \gamma = 1 = \text{area of original red section in Figure 1.2} \tag{1.4}$$

But the area of the square is also 1 (since its side has length 1), so the area of the crescent and the square are the same!

I don't know about you, but I think this is *really cool*.

1.5 Archimedes

Archimedes is arguably one of the best mathematicians of all time. He discovered many important things. One was determining the size of π. Archimedes calculated π by the exhaustion method (now called the "squeeze method" by some for the methodology used).

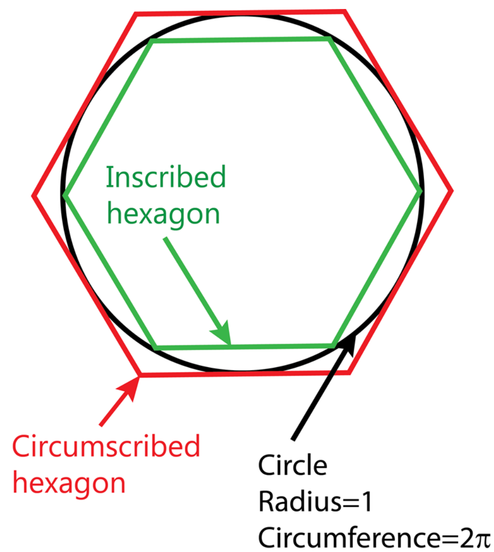

Figure 1.4: Schematic of the Archimedes exhaustion principle. The value of 2π lies in between the perimeter of the circumscribed polygon and the inscribed polygon. As the number of sides is made larger and larger the two perimeters approach each other, and the value of 2π.

Here is how the argument goes. Take a circle of unit radius. We will calculate the perimeter of an n-gon and compare it to the perimeter of a $2n$-gon both *inscribed inside* the circle.

The sides of the $2n$-gon and n-gon (see Fig. 1.5) can be written $s_{2n} = BD$, and $s_n = BC$. We can see that $\triangle ACP$, $\triangle ADB$, and $\triangle BDP$ are similar (this is because they are all right triangles with an acute angle of θ). One verifies this by comparing angles: we have $\angle ABC = 90^{\circ} - 2\theta = \angle ABP$, $\angle ABD = 90^{\circ} - \theta$, and $\angle PBD = \theta$. This follows because for any point on a circle the angle created from two diameter endpoints to the point is always $90°$, (sums of squares of chords)^2 = (diameter)^2. The proof of this is a homework problem.

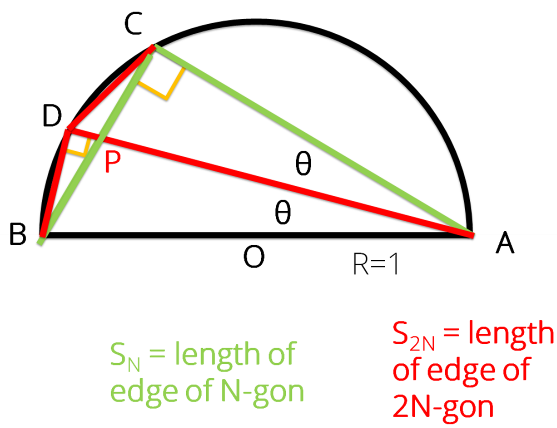

Figure 1.5: Triangles used to relate the side of one edge of the n-gon, given by $s_n = BC$ and the side of the $2n$-gon, given by $s_{2n} = BD$. The strategy is to first identify the similar triangles $\triangle ACP$, $\triangle ADB$, and $\triangle BDP$.

Observing these similar triangles (see also Fig. 1.6), we see

$$\frac{AB}{AD} = \frac{BP}{BD} \tag{1.5}$$

and hence

$$\frac{AC}{PC} = \frac{AD}{BD} \implies \frac{AC}{AD} = \frac{PC}{BD}. \tag{1.6}$$

Adding the top equation to the right part of the second equation yields

$$\frac{AB + AC}{AD} = \frac{BP + PC}{BD} = \frac{BC}{BD}. \tag{1.7}$$

Cross multiply

$$\frac{AB + AC}{BC} = \frac{AD}{BD}. \tag{1.8}$$

Squaring this equation gives

$$\frac{(AB)^2 + 2(AB)(AC) + (AC)^2}{(BC)^2} = \frac{(AD)^2}{(BD)^2} \tag{1.9}$$

Now we add 1 to both sides:

$$\frac{(AB)^2 + 2(AB)(AC) + (\mathbf{AC})^{\mathbf{2}} + (\mathbf{BC})^{\mathbf{2}}}{(BC)^2} = \frac{(\mathbf{AD})^{\mathbf{2}} + (\mathbf{BD})^{\mathbf{2}}}{(BD)^2} \tag{1.10}$$

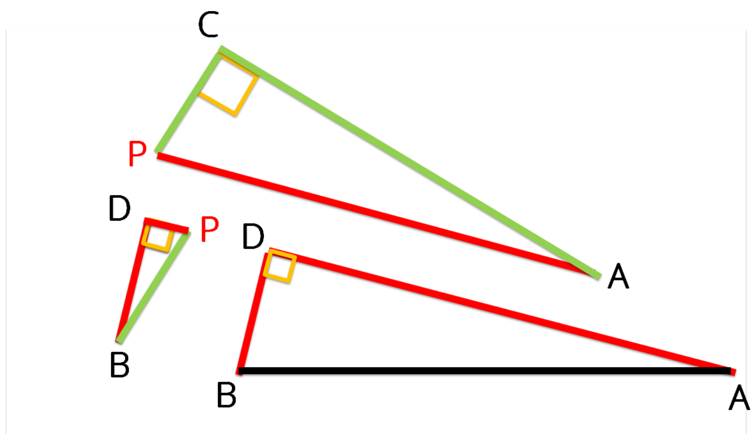

Figure 1.6: Closeup of the three triangles used in the proof. Matching the ratios of the lengths similar sides is used to finish the proof.

Note that each of the bold face terms are equal to $(AB)^2$. Substituting that result into the equation (in two places), then yields

$$\frac{2(AB)(AB+AC)}{(BC)^2} = \frac{(AB)^2}{(BD)^2} \tag{1.11}$$

We remember that $AB = 2 =$ diameter, $BD = s_{2n}$, and $BC = s_n$. We have

$$(AC)^2 = (AB)^2 - s_n^2 = 4 - s_n^2 \tag{1.12}$$

So

$$\frac{4(2+\sqrt{4-s_n^2})}{s_n^2} = \frac{4}{s_{2n}^2} \tag{1.13}$$

$$s_{2n}^2 = \frac{s_n^2}{2+\sqrt{4-s_n^2}}. \tag{1.14}$$

To start the recursion, We consider a hexagon, $n = 6$. The hexagon is made of six equilateral triangles, each with an edge equal to the radius of the circle, which is 1. So the initial perimeter is 6. (Be sure to draw picture and verify this is so.)

We solve the recursion numerically for the circumference $n \times s_n \rightarrow 2\pi r = 2\pi$ as n gets large. The results are given in the table (for the perimeter divided by 2).

N	S_N	Pi	T_N	Error width
6	3.0	3.14159265359	3.46410161514	0.464101615138
12	3.10582854123	3.14159265359	3.21539030917	0.109561767943
24	3.13262861328	3.14159265359	3.1596599421	0.0270313288163
48	3.13935020305	3.14159265359	3.14608621513	0.00673601208453
96	3.14103195089	3.14159265359	3.14271459965	0.0016826487548
192	3.14145247229	3.14159265359	3.14187304998	0.000420577694665
24576	3.14159264503	3.14159265359	3.14159267174	2.67075961347e-08

Figure 1.7: Table of the recursion results for increasing n. One can clearly see the convergence to π. This table is showing the perimeter divided by 2 for the inscribed and circumscribed polygons.

Archimedes also found the outer polygons to bound the value of π. The relationship between the n-gon and the $2n$-gon is

$$t_{2n} = \frac{2\sqrt{4+t_n^2}-4}{t_n} \tag{1.15}$$

Proving this is true will also be a homework problem.

Chapter 2

Everything you want to know about series but were afraid to ask

2.1 The geometric series

The book by Toeplitz discusses Zeno's paradox — that one can never go from point A to point B because it takes an infinite number of steps, if one takes half as big a step each time. Since one needs an infinite number of steps, one can never make it. But the flaw in Zeno's argument is that the sum of all of the steps is *finite*, which is why we can do it. We will prove

$$1+\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\cdots=2 \tag{2.1}$$

next.

We perform the proof in full generality, using an abstract x in the *geometric series*. Our goal is to show that

$$1+x+x^2+x^3+\cdots=\frac{1}{1-x} \tag{2.2}$$

for $|x|<1$. The strategy is to use the multiply by one trick. We have

$$\begin{aligned}&\left(1+x+x^2+x^3+x^4+\cdots+x^N\right)\frac{(1-x)}{(1-x)}\\&\quad=\left(1-x+x-x^2+x^2-\cdots-x^N+x^N-x^{N+1}\right)\frac{1}{(1-x)}.\end{aligned} \tag{2.3}$$

Note how all the terms except the first and the last have pairs appearing—one with a plus sign (black) and one with a minus sign (red)—so they will cancel. This shows that

$$1 + x + x^2 + x^3 + x^4 + \cdots + x^N = \frac{1 - x^{N+1}}{1 - x}. \tag{2.4}$$

If $|x| < 1$, then as $N \to \infty$, one will have $|x|^{N+1} \to 0$ (be sure you understand why).

So we take the limit $N \to \infty$ to obtain

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1 - x} \tag{2.5}$$

for $|x| < 1$. This sum is called the geometric series. It is an important result to remember (both the final answer and the methodology used to derive it). It will come up again many times.

2.2 The (dreaded) Taylor series

Finally, we talk about a MacLaurin and Taylor series. These are essentially the same thing. But to be completely precise, we note that the MacLaurin series is a Taylor series expanded about the point $x_0 = 0$. The Taylor series can be expanded about any point x_0. It is important to remember this, because most examples are MacLaurin series, but there are times when a Taylor series about a different point is needed.

Suppose we want a polynomial that has the same function value and same first n derivatives of a function at the origin. How do we make such a thing? Obviously the first term is the value of the function at the origin, or f(0). The second term must have the correct slope, so the coefficient of x is $\frac{df(0)}{dx}$, which we denote as $f^{(1)}(0)$, with the superscript indicating the number of derivatives. The third term is proportional to x^2. Since the second derivative of x^2 is 2, we must have its term be $f^{(2)}(0)$ multiplied by $\frac{x^2}{2}$. Hence, we have already found that the quadratic polynomial that satisfies this criterion is

$$f(0) + f^{(1)}(0)x + \frac{1}{2}f^{(2)}(0)x^2. \tag{2.6}$$

To get the full series, we just continue in the same fashion, recalling that

$$\frac{d^n}{dx^n}x^n = n! \tag{2.7}$$

yields

$$f(x) = \sum_{n=0}^{\infty} f^{(n)}(0)\frac{x^n}{n!}, \tag{2.8}$$

which, one can see, includes the quadratic polynomial we derived above.

As a check in your understanding, you should be able to take the nth derivative of the series, and evaluate it at $x = 0$, and show it is equal to $f^{(n)}(0)$. You must assume that you an interchange the order of taking a derivative and performing the sum when you carry out this calculation. Note that the result in Eq. (2.7) holds at every point x_0! This is why the form for the Taylor series will be exactly the same as the MacLaurin series (with the only change being $x \to x - x_0$).

Now we are ready for a worked-out example. Evaluate the MacLaurin series of $\sqrt{1+x}$. We need to calculate a number of different derivatives and evaluate them at $x = 0$.

$$\frac{d}{dx}\sqrt{1+x} = \frac{1}{2}\frac{1}{\sqrt{1+x}}\bigg|_{x=0} = \frac{1}{2} \tag{2.9}$$

$$\frac{d^2}{dx^2}\sqrt{1+x} = \frac{1}{2}\left(-\frac{1}{2}\right)\frac{1}{(1+x)^{\frac{3}{2}}}\bigg|_{x=0} = -\frac{1}{4} \tag{2.10}$$

$$\frac{d^3}{dx^3}\sqrt{1+x} = \frac{1}{2}\left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)\frac{1}{(1+x)^{\frac{5}{2}}}\bigg|_{x=0} = \frac{3}{8} \tag{2.11}$$

At this stage, it is common for students, who are “technicians,” to try to recognize the pattern and simply “hope” it continues that way forever. But it is better to carefully verify (via an induction argument), as a “practitioner” would, that it really holds. We will not go through these details here. Instead, we just show what we explicitly derived:

$$\sqrt{1+x} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3 + \dots \tag{2.12}$$

which follows from

$$\sqrt{1+x} = 1 + \frac{1}{2}x - \frac{1}{4}\left(\frac{x^2}{2}\right) + \frac{3}{8}\left(\frac{x^3}{6}\right) + \dots \tag{2.13}$$

You should try to do the full inductive argument to determine the general series on your own if you have never done this before.

2.3 Hyperbolic functions

This material does not exactly fit here, but it is useful for you to have a quick review of hyperbolic functions, which I have seen cause more than their fair share of consternation.

We begin with some basic definitions. The hyperbolic cosine is defined as

$$\cosh(x) = \frac{e^x + e^{-x}}{2} \tag{2.14}$$

and it is an *even* function of x. The hyperbolic sine is defined by

$$\sinh(x) = \frac{e^x - e^{-x}}{2} \tag{2.15}$$

and it is an *odd* function of x. The derivatives follow immediately from the definitions:

$$\frac{d\cosh(x)}{dx} = \frac{d}{dx}\frac{e^x + e^{-x}}{2} = \frac{e^x - e^{-x}}{2} = \sinh(x) \tag{2.16}$$

and

$$\frac{d\sinh(x)}{dx} = \frac{d}{dx}\frac{e^x - e^{-x}}{2} = \frac{e^x + e^{-x}}{2} = \cosh(x). \tag{2.17}$$

Recall how similar this is to the conventional trig functions except for some signs. Indeed, this is common with hyperbolics. Another example is the following identity

$$\cosh^2(x) - \sinh^2(x) = \frac{1}{4}(e^{2x} + 2 + e^{-2x}) - \frac{1}{4}(e^{2x} - 2 + e^{-2x}) = 1. \tag{2.18}$$

In summary, we have shown that

$$\frac{d\cosh(x)}{dx} = \sinh(x), \tag{2.19}$$

$$\frac{d\sinh(x)}{dx} = \cosh(x), \tag{2.20}$$

and

$$\cosh^2(x) = 1 + \sinh^2(x). \tag{2.21}$$

Plots of the hyperbolic functions are shown in Fig. 2.1.

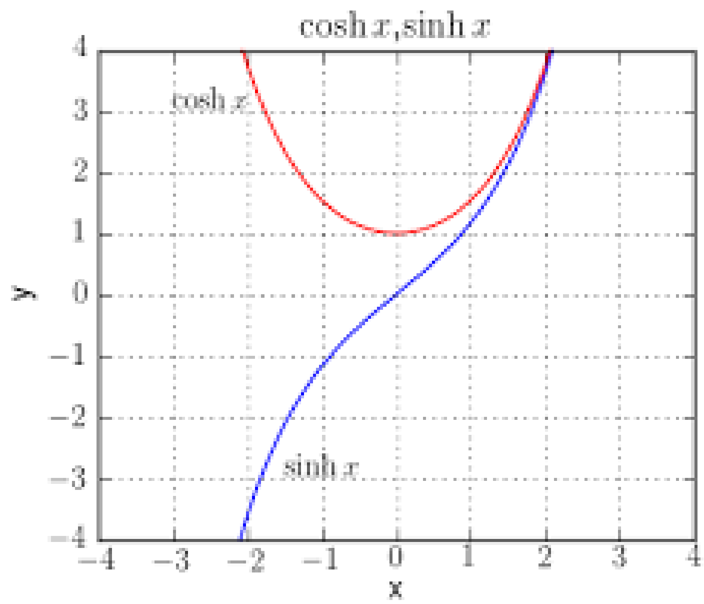

Figure 2.1: Plot of hyperbolic cosine (red) and sine (blue). One can immediately see their even and odd characters. The hyperbolic cosine is always larger than 1 and always large than the hyperbolic sine.

Chapter 3

Integrals and Limits

3.1 Origins of the concept of integration

We begin with the early work by the pioneers Archimedes and Cavalieri who showed how to integrate x^2 and x^k respectively. In modern formulas, we want to show how to integrate $\int_0^1 x^k$. The strategy is to relate these integrals to sums of powers of integers. We start by taking the interval of 0 to 1 and divide into n intervals running from $\frac{0}{n}$ to $\frac{n}{n}$.

The red rectangles have an area less than the integral, while the green rectangles cover a larger area. This sounds like we will be invoking the "squeeze principle" again. Indeed, we will. Let t_n denote the sum of the green rectangles and s_n the sum of the blue rectangles.

$$t_n = \left(\frac{1}{n}\right)^k \cdot \frac{1}{n} + \left(\frac{2}{n}\right)^k \cdot \frac{1}{n} + \left(\frac{3}{n}\right)^k \cdot \frac{1}{n} + ... + \left(\frac{n}{n}\right)^k \cdot \frac{1}{n} \tag{3.1}$$

$$s_n = \left(\frac{0}{n}\right)^k \cdot \frac{1}{n} + \left(\frac{1}{n}\right)^k \cdot \frac{1}{n} + \left(\frac{2}{n}\right)^k \cdot \frac{1}{n} + ... + \left(\frac{n-1}{n}\right)^k \cdot \frac{1}{n} \tag{3.2}$$

So

$$s_n \leq \int_0^1 x^k \leq t_n \tag{3.3}$$

But

$$t_n = \sum_{j=1}^{n} \frac{j^k}{n^{k+1}}, \tag{3.4}$$

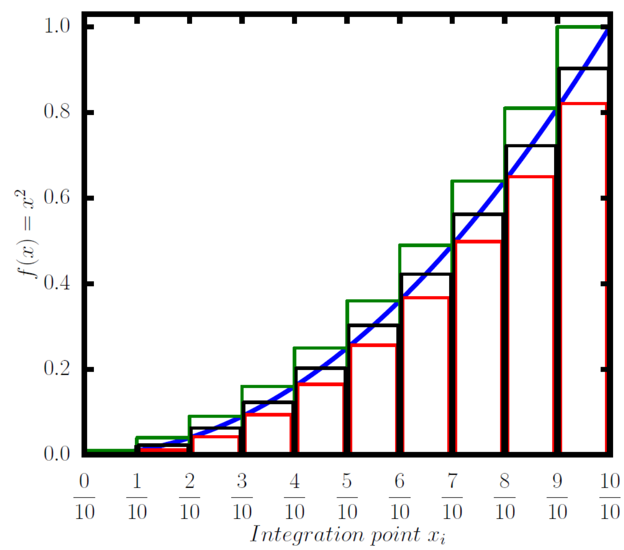

Figure 3.1: Three different rectangular integration routines for $\int_0^1 x^k\,dx$ with $k = 2$. The function is plotted in blue, the three different numerical integration schemes are left point (red), midpoint (black) and right point (green). One can clearly see the integral is bounded to lie in between the sum of the red and green rectangular areas.

and

$$s_n = \sum_{j=0}^{n-1} \frac{j^k}{n^{k+1}} = \sum_{j=1}^{n-1} \frac{j^k}{n^{k+1}}, \tag{3.5}$$

where in the second equality in the s_n equation, we note that we can drop the $j = 0$ term.

Now we invoke the critical piece of the squeeze argument. To begin, note that $t_n - s_n = \frac{1}{n}$ so as $n \to \infty$, $t_n - s_n \to 0$ or $t_n \to s_n$ and this limit is the integral. How do we know this result, that $t_n - s_n = \frac{1}{n}$? Each term in s_n also appears in t_n. But t_n has one more term—its last term. Hence, the difference is precisely that last term $\frac{n^k}{n^{k+1}} = \frac{1}{n}$. Then the rest of the argument follows as above. To determine the final result, we need to now evaluate these finite sums of powers of integers.

3.2 Sums of powers of integers

The skill to learn how to sum powers of integers is a useful one. You might think it is an odd thing to do, as we cannot extend these sums to infinity (unlike some popular youtube videos would say), because such sums always *diverge*. But we can actually get closed-form expressions for *finite summations*. And this is a rather marvelous result. We show you how to do this next.

Define

$$\mathrm{sum}_k(n) = \sum_{j=1}^{n} j^k = \text{ sum of first n integers raised to the k power.} \tag{3.6}$$

The simplest case is $k = 1$.

$$\mathrm{sum}_1(n) = 1 + 2 + 3 + ... + n. \tag{3.7}$$

We can write $\mathrm{sum}_1(n)$ in reverse order underneath, and add down the columns:

$$\begin{aligned} \mathrm{sum}_1(n) &= 1 + \quad 2 \quad + \quad 3 \quad + \ldots + n \\ \mathrm{sum}_1(n) &= n + (n-1) + (n-2) + \ldots + 1 \end{aligned} \tag{3.8}$$

$$2 \times \mathrm{sum}_1(n) = \underbrace{(n+1) + (n+1) + (n+1) + ... + (n+1)}_{n \text{ terms}} \tag{3.9}$$

There are n terms of $(n + 1)$, so

$$\mathrm{sum}_1(n) = \frac{n(n+1)}{2}. \tag{3.10}$$

This result is very direct, simple and neat.

Unfortunately, this approach is not easily generalized to other k powers. So we have to proceed a different way. We do so by computing differences of finite sums in two different ways. Consider

$$\sum_{j=1}^{n} \left[(j+1)^2 - j^2\right] = (n+1)^2 - 1 = n^2 + 2n, \tag{3.11}$$

which follows since many of the terms (all except the first term of the second sum and the last term of the first sum) cancel when the sums are subtracted from ach other. Now, we compute a second way, by expanding the terms in the sums (this is mathematically fine because all sums are *finite*). We find

$$\sum_{j=1}^{n} \left[(j+1)^2 - j^2\right] = \sum_{j=1}^{n} \left(\cancel{j^2} + 2j + 1 - \cancel{j^2}\right) = \sum_{j=1}^{n} 2j + \sum_{j=1}^{n} 1 \tag{3.12}$$

We realize that

$$\sum_{j=1}^{n} 2j = 2\sum_{j=1}^{n} 1 \tag{3.13}$$

and

$$\sum_{j=1}^{n} 1 = n. \tag{3.14}$$

So we immediately discover (after equating the two ways to evaluate the summations) that

$$2\,\mathrm{sum}_1(n) = n^2 + 2n - n = n(n+1) \tag{3.15}$$

so

$$\mathrm{sum}_1(n) = \frac{n(n+1)}{2} \tag{3.16}$$

as before.

But this strategy can be generalized to higher powers of k.

We can check the next case where we work with cubes instead of squares. We have

$$\sum_{j=1}^{n} \left[(j+1)^3 - j^3\right] = (n+1)^3 - 1 = n^3 + 3n^2 + 3n \tag{3.17}$$

and

$$\sum_{j=1}^{n} \left[(j+1)^3 - j^3\right] = \sum_{j=1}^{n} \left(\cancel{j^3} + 3j^2 + 3j + 1 - \cancel{j^3}\right) \tag{3.18}$$

$$= 3\,\mathrm{sum}_2(n) + 3\,\mathrm{sum}_1(n) + \mathrm{sum}_0(n) \tag{3.19}$$

$$= n^3 + 3n^2 + 3n. \tag{3.20}$$

So

$$\begin{aligned}
\mathrm{sum}_2(n) &= \frac{1}{3}\left(\underbrace{n^3+3n^2+3n}_{\text{sum first way}}\underbrace{-\frac{3}{2}n^2-\frac{3}{2}n}_{-3\,\mathrm{sum}_1(n)}\underbrace{-n}_{-\mathrm{sum}_0(n)}\right) && (3.21)\\
&= \frac{1}{3}\left(n^3+n^2+\frac{1}{2}n\right) && (3.22)\\
&= \frac{1}{6}n\left(2n^2+3n+1\right) && (3.23)\\
&= \frac{1}{6}n\left(2n+1\right)\left(n+1\right). && (3.24)
\end{aligned}$$

Hence, we have found that

$$\sum_{j=1}^{n} j^2 = \frac{1}{6}n(2n+1)(n+1). \tag{3.25}$$

For the $k=2$ case, the integral of the parabola then gives

$$\int_0^1 x^2\,dx = \lim_{n\to\infty}\frac{1}{6}n(2n+1)(n+1)\frac{1}{n^3} = \frac{1}{3} \tag{3.26}$$

as expected. One can use this technique to calculate higher integer powers as well. Cavalieri extended this to all the way to $k=10$. The strategy is to extend the expression for the difference of the sums of powers to higher k values. The highest-order terms, proportional to j^k cancel in the two summations. Then the procedure continues as we did above. We need to use the results for all previous powers to obtain the final answer. Note that there does not seem to be any simply pattern to these final answers, so we do not try to establish them by induction. But if you like, you could determine the result for the highest power of n (proportional to n^{k+1}), which is what is needed for the calculation of the integral.

Fermat used a slightly different method, with an infinite number of steps and the identity

$$\lim_{x\to 1}\frac{1-x^{k+1}}{1-x} = \lim_{x\to 1}\left(1+x+x^2+x^3+...+x^k\right) = k+1 \tag{3.27}$$

to prove the integral result for all integer k.

3.3 Issues with defining an integral

The book discusses issues with defining an integral in greater depth, but the main results are

- The integral is well defined for piecewise monotonic (strictly increasing or decreasing) functions
- Strange functions (like the one in section 17) show that one has to be careful and precise in defining the integral. This often does not play a role in integrals that arise in physics, but occasionally is relevant in some areas like Cantor sets or strange attractors. The field of real analysis spends significant time sorting out all possible subtleties in how one defines an integral.
- *The definite integral is not the area under the curve.* It is often that one can think of it as the *signed* area under the curve, where negative values correspond to negative areas.

3.4 Numerical integration

We do not go into detail into numerical techniques in this class, but it is important for you to know some of the basics. Below is a crash course on numerical quadrature.

While the definite integral is defined in the limit where the maximal step size approaches 0, we must work with a finite step size to calculate a numerical approximation to the integral. The simplest quadrature rule is a left point, midpoint, or right point integration rule. These are demonstrated in Fig. 3.2.

We call the coordinate where the function is evaluated x_i, and the value of the function used $f(x_i)$. Then the numerical approximation to the integral replaces the integral with a finite sum of terms

$$\int_a^b f(x)\,dx = \Delta x \sum_{i=1}^{n} f(x_i) \tag{3.28}$$

where the $\{x_i\}$ are chosen as described in Fig. 3.2, corresponding to the specific integration rule that is being used. For example, in left-point integration, we choose $\Delta x = (b-a)/n$ and $x_i^{\text{left}} = (i-1)\Delta x$. The right point rule is

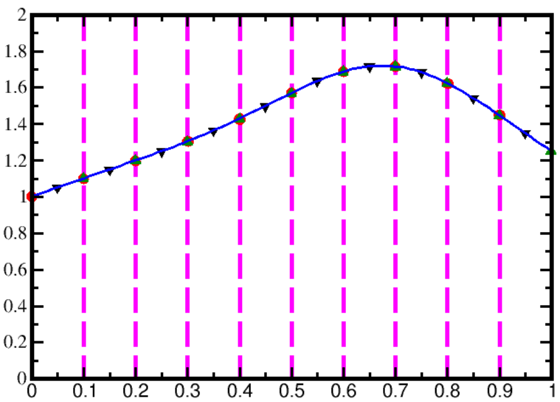

Figure 3.2: Different rectangular integration rules. In this example, we have ten rectangles whose area we sum to approximate the integral $\int_0^1 f(x)\,dx$ ($f(x)$ in blue). The left-point rule, takes the values of the functions at the red circles for each rectangle. The right point rule takes the values with the green triangles, while the midpoint rule takes the black triangular values. In the limit as the step size goes to zero, all rules will give the same answer, but their results differ for $\Delta x \neq 0$.

$x_i^{\text{right}} = i\Delta x$. The midpoint rule, averages the two and is $x_i^{\text{mid}} = \left(i - \frac{1}{2}\right)\Delta x$. In all cases $1 \leq i \leq n$ ($n = 10$ in the figure). These first rules are rectangular integration rules. Images most effectively illustrate the different methods, as can be seen in the corresponding Figs. 3.3-3.5.

The trapezoidal rule uses trapezoids to fit the curve (red trapezoids). This result is exactly equivalent to a different set of rectangles centered at the gridpoints; in this case, for the first and last points, only one-half of the corresponding rectangle contributes.

The next to consider is the so-called Simpson's rule, which approximates the integral via summing with weights that exactly integrate constant, linear, and quadratic functions. It can also be thought of as approximating the function with a quadratic for every sequence of three adjacent points on the grid. We will not draw a figure for this, but you can certainly imagine what it looks like. As with the trapezoidal rule, the construction simplifies, and can be thought of as integration over rectangles with the weights alternating

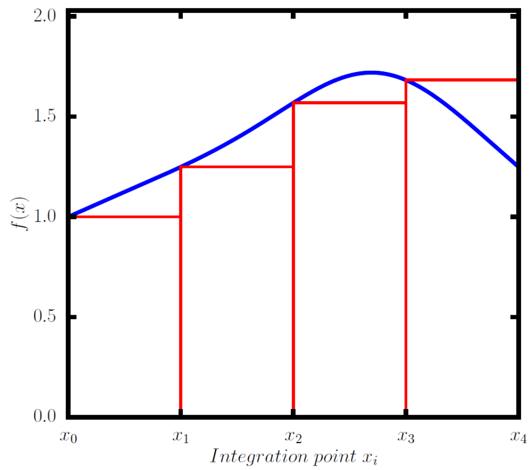

Figure 3.3: Left-hand integration rule with four points, given by $\int_{x_0}^{x_4} f(x)dx \approx \Delta x\,[f(x_0) + f(x_1) + f(x_2) + f(x_3)]$.

from 4/3 to 2/3 with the endpoints given by 1/3 again:

$$\int_{x_0}^{x_5} f(x)dx \approx \Delta x \left[\frac{1}{3}f(x_0) + \frac{4}{3}f(x_1) + \frac{2}{3}f(x_2) + \frac{4}{3}f(x_3) + \frac{1}{3}f(x_4)\right] \quad (3.29)$$

The sum alternates $\frac{4}{3}$ weight and $\frac{2}{3}$ weight, with $\frac{1}{3}$ weight on the endpoints. The generalization of Simpson's rule to higher polynomials is called Romberg integration. Many prefer this type of extrapolation technique to other methods. One can think of it as a technique that tries to extrapolate all the way to $\Delta x \to 0$.

There is one other technique for integration called Gaussian integration where the abscissae and weights are determined from a prescribed formula that exactly integrates a weight function times a polynomial of some degree. In this case, the spacing of the grid points is not uniform and it changes for every point as the number of points in the sum changes. This can make it inconvenient for computation and determining accuracy when one computes for different numbers of grid points. The reason why, is for grids that are evenly spaced, we simply double the number of points. Then half of the old grid points are on the new grid and we do not need to recalculate the function at those points. This never holds with Gaussian integration techniques. They are used, however, because if you have an integral of the form of a Gaussian

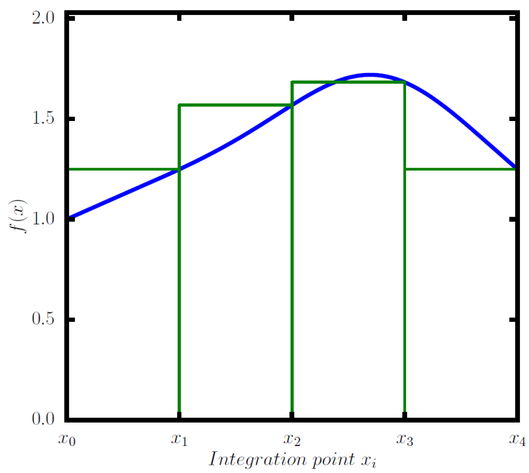

Figure 3.4: Right-hand integration rule with four points, given by $\int_{x_0}^{x_4} f(x)dx \approx \Delta x\,[f(x_1) + f(x_2) + f(x_3) + f(x_4)]$.

integration weight function, this tailor-made approach is likely to be superior. Calculating the weights and the gridpoints, however, is complicated.

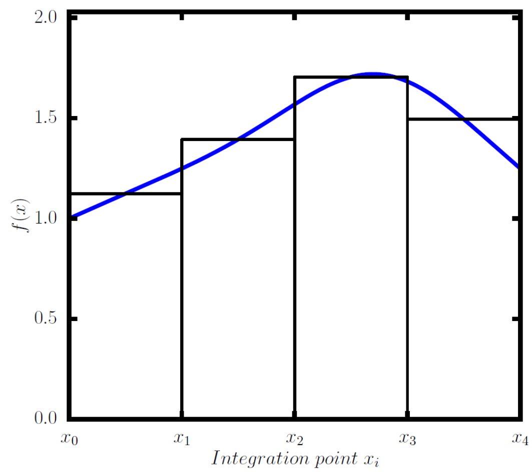

Figure 3.5: Midpoint integration rule with four points, given by $\int_{x_0}^{x_4} f(x)dx \approx \Delta x \left[f\left(\frac{x_0+x_1}{2}\right) + f\left(\frac{x_1+x_2}{2}\right) + f\left(\frac{x_2+x_3}{2}\right) + f\left(\frac{x_3+x_4}{2}\right)\right]$.

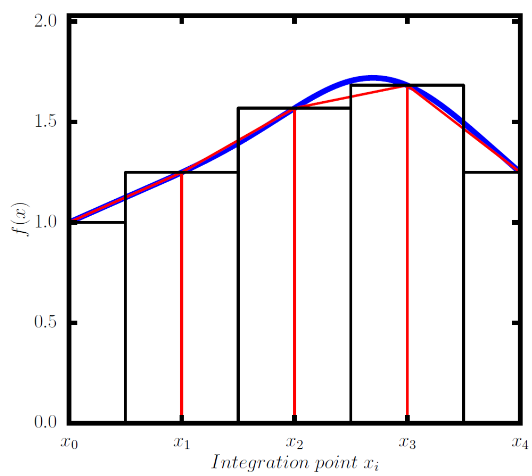

Figure 3.6: Trapezoidal integration rule with four points. The red trapezoids show the integration following the direct rule. It is equivalent to the black rectangles with the first and last counted half. The rule is given by $\int_{x_0}^{x_4} f(x)dx \approx \Delta x \left[\frac{1}{2}f(x_0) + f(x_1) + f(x_2) + f(x_3) + \frac{1}{2}f(x_4)\right]$.

Chapter 4

Tangents and Logarithms

4.1 Optimizations with constraints

The book by Toeplitz describes how a significant effort was expended by mathematicians on finding tangents to curves and how calculus makes it much easier to do. We will examine, as an example, one such derivative problem—for the set of rectangles of fixed perimeter p, what shape has the largest area?

Recall that area = length×width, and perimeter = 2×(length + width).

So if we let w denote the width and p denote the perimeter, then we have

$$l = \text{length} = \frac{p}{2} - w. \tag{4.1}$$

The area satisfies

$$A = l \times w = \left(\frac{p}{2} - w\right) \times w = \left(\frac{p}{2}\right) w - w^2. \tag{4.2}$$

Differentiating to determine the maximum gives

$$\frac{dA}{dw} = \frac{p}{2} - 2w = 0 \to w = \frac{p}{4}. \tag{4.3}$$

Solving for the length, then yields

$$l = \frac{p}{2} - \frac{p}{4} = \frac{p}{4} = w \to l = w, \tag{4.4}$$

which means the shape is a square (because the length is equal to the width)!

Some people would like to solve this problem in a much simpler way by saying such a rectangle must be a special rectangle. But the only special rectangle we know is a square. So it must be square. Sometimes such arguments are deep and meaningful because the arguments are based on symmetry principles. Other times, it is just good luck that it gives the right answer. It is safer, at this stage of your career, to err on the side of producing a rigorous argument to support such statements, rather than falling back on a "symmetry" or "unique" argument.

4.2 Trigonometric tables

I want to take the remainder of our time discussing sine and logarithm tables. Back in the period from 200 BC to the early to mid 1900's, nearly all calculations were done with tables or instruments that acted as tables, such as slide rules.

The Greeks constructed sine tables, while logarithms didn't come about until the early 1600's. The accuracy needed was one part in ten million, or 7 digits, in order to perform astronomical calculations accurately enough. In other words, this is what Kepler needed to establish his laws of planetary motion. Trying to develop such tables was very demanding but it was a key to technological advance at the time.

The generation of a sine table was known to the Greeks. A 30–60–90 triangle gives $\sin(30°)$. The construction of a regular pentagon gives $\sin(36°)$. Archimedes formulas, which can be employed to produce $\sin\left(\frac{x}{2}\right)$ from $\sin(x)$, give us $\sin(15°)$ and $\sin(18°)$. The Greeks also knew sine addition formulas such as

$$\sin(\alpha \pm \beta) = \sin(\alpha)\cos(\beta) \pm \cos(\alpha)\sin(\beta). \tag{4.5}$$

So using this, one obtains $\sin(18° - 15°) = \sin(3°)$. Archimedes again gives $\sin(1\frac{1}{2}°)$ and $\sin(\frac{3}{4}°)$. Then, they use the identity

$$\frac{\sin(x)}{\sin(y)} < \frac{x}{y} \text{ for } 0 < y < x < 90°, \tag{4.6}$$

but recall that the angles must be expressed in radians for all of these calculations.

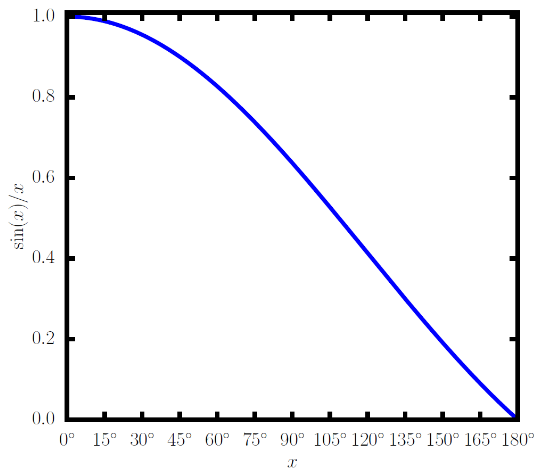

Figure 4.1: The function $\sin(x)/x$ for the range from 0 to π. Note how it is always a decreasing function of x. Such a function is called monotonic.

The identity follows from $\frac{\sin(x)}{x} < 1$ being a monotonic decreasing function of x (see Fig. 4.1) so

$$\frac{\sin(x)}{x} < \frac{\sin(y)}{y} \text{ for } y < x. \tag{4.7}$$

Cross multiplying gives us

$$\frac{\sin(x)}{\sin(y)} < \frac{x}{y} \text{ for } y < x. \tag{4.8}$$

This general result implies that

$$\frac{\sin(1\frac{1}{2}^\circ)}{\sin(1^\circ)} < \frac{3}{2} \text{ or } \frac{2}{3}\sin\left(1\frac{1}{2}^\circ\right) < \sin(1^\circ) \tag{4.9}$$

and

$$\frac{\sin(1^\circ)}{\sin(\frac{3}{4}^\circ)} < \frac{4}{3} \text{ or } \sin(1^\circ) < \frac{4}{3}\sin\left(\frac{3}{4}^\circ\right). \tag{4.10}$$

Using the values for $\sin(1\frac{1}{2}^\circ)$ and $\sin(\frac{3}{4}^\circ)$ pins (or better squeezes) the value of $\sin(1^\circ)$ to 4-5 digits of accuracy. They then computed $\sin(\frac{1}{2}^\circ)$ from Archimedes and finally used the addition formula to generate sine tables for every $\frac{1}{2}^\circ$.

4.3 Tables of logarithms

Log tables were even more valuable. Multiplying two 7-digit numbers to keep 7 digits of accuracy was tedious. But using a logarithm table reduced multiplication to addition because $\ln(xy) = \ln(x)+\ln(y)$, or calculating roots to division because $\ln(\sqrt[n]{x}) = \frac{1}{n}\ln(x)$.

We start with an example. Compute $\sqrt[3]{36000}$. A log table with a specific step size tells us that

12809	35996.4763
12810	36000.0759

where the first entry is the step (or exponent) and the second is $(1+x)^{step}$, where x is the value (step size) used in generating the table.

To solve this problem, we first interpolate to find $\log(3600) = 12809.98$. We then divide by 3 to get 4299.99. We next go to the table and find the step associated with this number (not shown here) to find the cube root (ans: 33.019272). So, one could compute complex things with these tables.

Generating these tables was mind numbing because it had to be done by hand. A four digit accuracy table required 2,300 steps. So, here are the first few steps of an example table with $x = 0.0001$.

step	$10,000 \times (1+\frac{1}{10,000})^{step}$
0	10,000.0000
1	10,001.0000
2	10,002.0001
3	10,003.0003

The key trick to making the table is that one never uses multiplication. They are formed instead *by addition* using the following result:

$$a \times \left(1 + \frac{1}{10,000}\right) = a + \frac{a}{10,000}. \tag{4.11}$$

The addition is done in a simple fashion. Write down the previous number a. Take the same number and shift it four digits to the right. Then add them together, truncating terms that are beyond the desired accuracy. Then one repeats. Again. And again. And again.

Generating these tables in base 10 seemed to be most convenient for calculations, but to get 7 digits of accuracy requires 23,000,000 steps and no

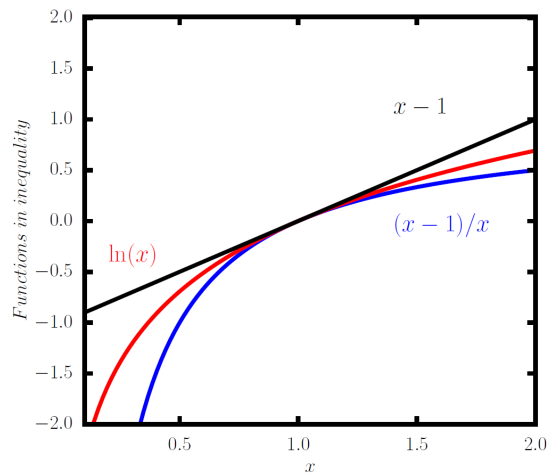

Figure 4.2: Functions used in the inequality. Blue is $\frac{x-1}{x}$, red is $\ln(x)$, and black is $x-1$. Note how they remain ordered even though they become equal at $x = 1$.

one could take up such work. This was further confounded by the fact that any error made at one step, invalidates all further entries in the table.

John Napier figured out some simplifications, that allowed one to actually construct such a table.

1. Compute a table with entries $(1 + \frac{1}{100})^{step}$ for each integer step.

2. Since steps in different tables are related by fixed ratios, compute every 100^{th} entry of a $(1 + \frac{1}{10000})^{step}$ table by multiplying the entries in the lower-precision table by that specific factor. Then fill in the first 100 entries of the higher precision table, and by simple addition, one finds all subsequent entries.

3. The remaining problem was the ratio between the low-precision and the high-precision table, which was not clear how to compute. Napier sought for an *absolute* logarithm table to resolve this issue.

4. To do so, Napier examined how the geometric table entries (called $f(x) = (1 + x)^{step}$) varied with respect to the table parameter x and found that
$$\lim_{x_1 \to x} \frac{f(x) - f(x_1)}{x - x_1} = \frac{c}{x} = f'(x). \tag{4.12}$$

So the simplest table would have $c = 1$, which defines the "absolute" logarithm table.

5. Napier further notices that if we compare the curves

$$\frac{x-1}{x}, \qquad \ln(x), \qquad x-1 \tag{4.13}$$

all three vanish at $x = 1$ (and hence meet each other) and the slopes are

$$\frac{1}{x^2}, \qquad \frac{1}{x}, \qquad 1 \tag{4.14}$$

respectively. Since $\frac{1}{x^2} < \frac{1}{x} < 1$, for $x > 1$ (and the opposite for $x < 1$), we find the curves remain strictly ordered with respect each other, as shown in Fig. 4.2. Hence,

$$\frac{x-1}{x}, \qquad \ln(x), \qquad x-1 \tag{4.15}$$

becomes, after dividing by $x - 1$,

$$\frac{1}{x} < \frac{\ln(x)}{x-1} < 1, \tag{4.16}$$

for all x. Now, let $x = \frac{a}{b} \rightarrow \frac{1}{a} < \frac{\ln(a)-\ln(b)}{a-b} < \frac{1}{b}$ for $a > b$. This identity allows the factor between different tables to be found and further allows one to compute the table entries more easily for the higher-precision tables.

What base did Napier use? Let $x = \frac{n+1}{n}$ in the inequality in Eq. (4.16). Then we find that

$$\frac{n}{n+1} < \frac{\ln(1+\frac{1}{n})}{\frac{1}{n}} < 1 \tag{4.17}$$

or

$$\frac{n}{n+1} < \ln\left(1+\frac{1}{n}\right)^n < 1. \tag{4.18}$$

now, as we take $n \rightarrow \infty$, we find that $(1+\frac{1}{n})^n \rightarrow e$, so $\ln(e) = 1 \rightarrow \text{base} = e$.

They then used very clever tricks, described at the end of section 22 of the Toeplitz book, to compute tables with seven digits of accuracy. These tables were extremely influential for hundreds of years and were needed for many different calculations. The first important problem was allowing Kepler to develop his rules of planetary motion.

Note that we hardly use such tables anymore. But it is important that how they were constructed not become a lost art. There is much insight to be found from understanding how one constructed these tables.

Chapter 5

Fundamental Theorem of Calculus and Manipulation of Integrals

5.1 Fundamental theorem of calculus

We start with the fundamental theorem of calculus: If $F(t) = \int_a^t f(x)\,dx$ with $a < t < b$ and if $f(x)$ is continuous and monotonic for $a \le x \le b$, then $F'(t) = f(t)$. (Barrow)

To prove the fundamental theorem, we just compute the derivative

$$\lim_{t_1 \to t} \frac{F(t_1) - F(t)}{t_1 - t} = \lim_{t_1 \to t} \frac{\int_0^{t_1} f(x)\,dx - \int_0^t f(x)\,dx}{t_1 - t} = \lim_{t_1 \to t} \frac{\int_t^{t_1} f(x)\,dx}{t_1 - t} \quad (5.1)$$

Since $f(t) < f(x) < f(t_1)$ for $t < x < t_1$ because f is monotonic, we have

$$(t_1 - t) f(t) < \int_t^{t_1} f(x)dx < (t_1 - t) f(t_1) \quad (5.2)$$

$$\implies f(t) < \frac{F(t_1) - F(t)}{t_1 - t} < f(t_1) \quad (5.3)$$

$$\implies \lim_{t_1 \to t} \frac{F(t_1) - F(t)}{t_1 - t} = f(t) \quad (5.4)$$

since f is continuous.

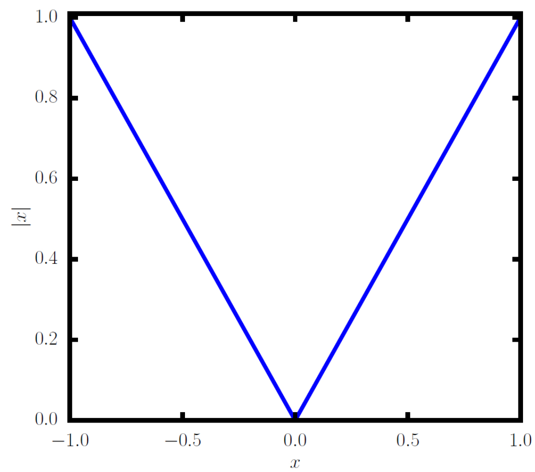

Figure 5.1: Absolute value function. Note how the derivative is -1 for $x < 1$ and 1 for $x > 1$. The limit as $x \to 0$ does not exists for the derivative because the left limit is not equal to the right limit.

The fundamental theorem of calculus has an obvious corollary as well. It is, in fact, the most common use of the theorem.

Corollary: If $\phi'(x) = f(x)$, $\phi(x) = F(x) + c$ where c is a constant.

If f is continuous at x, then the left and right limits match: $\lim_{x \to x_0^+} f(x) = \lim_{x \to x_0^-} f(x)$, but continuity also requires such a limit to exist in the first place (see figure 81 in the Toeplitz book for an example of a function with no limit). Note, however, that continuity does not imply that f is differentiable. $f(x) = |x|$ is a classic example: the function is continuous but is not differentiable at $x = 0$.

The converse is, however, true. If a function is differentiable at x, then it must also be continuous at x.

5.2 Product Rule

The product rule, sometimes called the Leibnitz rule, shows how one can calculate the derivative of a product of functions. The formula is well known to any student in a Calculus I class.

$$w(x) = u(x)v(x) \implies w'(x) = u'(x)v(x) + u(x)v'(x) \tag{5.5}$$

Proof:

$$w'(x) = \lim_{x_1 \to x} \frac{u(x_1)v(x_1) - u(x)v(x)}{x_1 - x} \tag{5.6}$$

$$= \lim_{x_1 \to x} \frac{u(x_1)v(x_1) - u(x)v(x_1) + u(x)v(x_1) - u(x)v(x)}{x_1 - x} \tag{5.7}$$

$$= \lim_{x_1 \to x} \frac{u(x_1) - u(x)}{x_1 - x} v(x_1) + u(x) \frac{v(x_1) - v(x)}{x_1 - x} \tag{5.8}$$

$$= u'(x)v(x) + u(x)v'(x) \tag{5.9}$$

Note the use of the "add zero" trick in the second line (terms in red). Now, since v is differentiable, it is also continuous, so we can replace $v(x_1)$ by $v(x)$ in the last line.

We will make use of the product rule to derive the important result called integration by parts.

5.3 Integration by Parts

Integating the Leibnitz rule (via the fundamental theorem of calculus) shows us that

$$u(x)v(x) = \int u'(y)v(y)\,dy + \int u(y)v'(y)\,dy. \tag{5.10}$$

Rearranging, we find

$$\int u'(y)v(y)\,dy = u(y)v(y) - \int u(y)v'(y)\,dx \tag{5.11}$$

Integration by parts is usually used for a definite integral, as follows:

$$\int_a^b u'(x)v(x)dx = u(x)v(x)\Big|_a^b - \int_a^b u(x)v'(x)dx. \tag{5.12}$$

Integration by parts is an extremely useful technique for evaluating integrals. For example, consider for $n \neq -1$: $\int x^n \ln(x)\,dx$. We take $u' = x^n$ and $v = \ln(x)$.

$$\int x^n \ln(x)\,dx = \frac{x^{n+1}}{n+1} \ln(x) - \int \frac{x^{n+1}}{n+1} \frac{1}{x}\,dx = \frac{x^{n+1}}{n+1} \ln(x) - \frac{x^{n+1}}{(n+1)^2} \tag{5.13}$$

In particular, for $n = 0$, we have $\int \ln(x)dx = x\ln(x) - x$ which would have been difficult to guess by any other means.

5.4 Inverse Chain Rule

The so-called chain rule for the derivatives of functions of functions is the following: if $g(u)$ is a differentiable function of u and $u(x)$ is a differentiable function of x, then

$$\frac{dg(u)}{dx} = \frac{dg(u)}{du}\frac{du}{dx}. \tag{5.14}$$

This is one of the most useful relations of differential calculus. It can simplify derivatives if you are clever. For example, consider $\frac{d}{d\theta}\left(\sin^2\theta + \frac{1}{\sin\theta}\right)$. It is easy to calculate as

$$\frac{d}{d\sin\theta}\left(\sin^2\theta + \frac{1}{\sin\theta}\right)\frac{d\sin\theta}{d\theta} = \left(2\sin\theta - \frac{1}{\sin^2\theta}\right)\cos\theta \tag{5.15}$$

In fact, we may be doing this subconsciously as we are following the rules for derivatives, but there are many situations where using the chain rule in this fashion can make calculations much easier to finish.

Related to this idea is the "inverse" of the chain rule

$$\int g(u)u'(x)\,dx = \int g(u)\frac{du}{dx}dx = \int g(u)du, \tag{5.16}$$

which is another valuable tool for integration. An example is $\int \frac{\ln x}{x}dx$. We have $g(u) = \ln x$ and $u'(x) = \frac{1}{x}$, so that

$$\int \frac{\ln x}{x}dx = \int uu'dx = \int udu = \frac{u^2}{2} = \frac{1}{2}(\ln x)^2. \tag{5.17}$$

Another example is

$$\int \frac{1}{x\ln x}dx = \int \frac{u'}{u}dx = \int \frac{du}{u} = \ln u = \ln(\ln x) \tag{5.18}$$

and so on.

5.5 Inverse Functions

Inverse functions are important throughout physics and math. One thing you must remember is that a function relates *one value* to each argument

x, which makes it single-valued: for each x, there is only one $f(x)$. If we want to compute the inverse of $f(x)$, for each y we must find the x such that $f(x) = y$, or $f^{-1}(y) = x$. We must have only one x value for each y value. This is, in general, where the complexity occurs, because the definition of a single-valued function does not guarantee that the inverse function is also single-valued. Indeed, many are not, which means we must restrict the domain for the inverse function to have it be single valued.

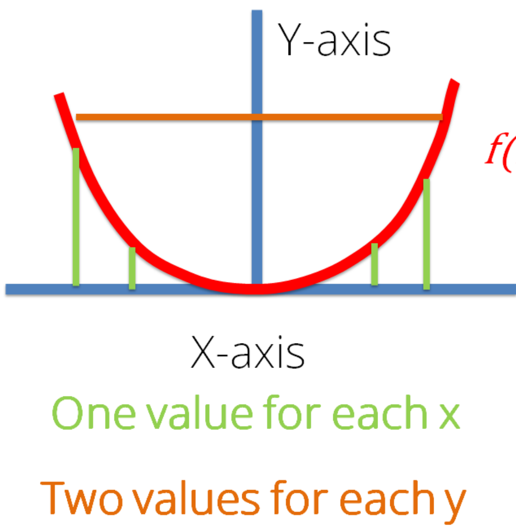

Figure 5.2: Schematic of functions and inverse functions for $y = x^2$.

Take a simple function like $f(x) = x^2$ for $-\infty < x < \infty$. Note how for every x, there is one and only one $y = x^2$ value (green lines in Fig. 5.2), so f is a single valued function. But for the inverse, if we set a value of y, there are two roots, one positive and one negative (red horizontal line in Fig.‘5.2), so there are two inverse functions defined on different domains: $f^{-1}(y) = \sqrt{y}$ gives answers in the range 0 to ∞ and $\tilde{f}^{-1}(y) = -\sqrt{y}$ gives answers in the range $-\infty$ to 0. Both are valid inverse functions. For general functions that are not strictly monotonic, the inverse functions will be defined on different ranges. This is particularly true for trigonometric functions which we will treat in the next lecture.

5.6 Examples

We end the chapter with some examples.
1. Differentiate a^x.

Solution: $a^x = e^{x \ln a}$

$$\frac{d}{dx} a^x = e^{x \ln a} \frac{d}{dx} (x \ln a) = e^{x \ln a} \ln a = a^x \ln a. \tag{5.19}$$

So $\frac{d}{dx} a^x = a^x \ln a$. Note how $\frac{d}{dx} e^x = e^x$ since $\ln e = 1$.

2. Compute the following integral:

$$\int \frac{x^4 - a^4}{x^2 + a^2} dx \tag{5.20}$$

Solution: It looks impossible to get a simple answer at first, but recall that

$$x^4 - a^4 = \left(x^2 - a^2\right)\left(x^2 + a^2\right) \tag{5.21}$$

So the integral is

$$\int \left(x^2 - a^2\right) dx = \frac{x^3}{3} - xa^2 \tag{5.22}$$

3. Compute the following integral:

$$\frac{d}{dx} \ln \left(\frac{x^2 + 1}{x^2 - 1}\right) \tag{5.23}$$

Solution:

$$\frac{d}{dx} \ln \left(\frac{x^2 + 1}{x^2 - 1}\right) = \frac{d}{dx} \ln \left(x^2 + 1\right) - \frac{d}{dx} \ln \left(x^2 - 1\right) = \frac{2x}{x^2 + 1} - \frac{2x}{x^2 - 1} = \frac{-4x}{x^4 - 1}. \tag{5.24}$$

Could you show or recognize that

$$\int \frac{x}{1 - x^4} dx = \frac{1}{4} \ln \left(\frac{x^2 + 1}{x^2 - 1}\right)? \tag{5.25}$$

To do this one would first expand by partial fractions, but one would then have to recognize that integrals of the form $\int \frac{2x}{x^2+1} dx$ are also of the form $\int \frac{du}{dx} dx = \int du = u$ for $u = \ln(x^2 + 1)$, which is often difficult to remember or recognize.