# No Bullshit Guide to Math and Physics

Preview and sample chapter

The full book has 562 pages and includes 55 exercises and 252 solved problems. This preview has been chosen to showcase some key aspects of the book (chapter intros, formulas, explanations, and figures). Buy the full book for only US$29 at https://gum.co/noBSmathphys .

No bullshit guide to math and physics
by Ivan Savov

Contents

Placement exam

The answers[1] to this placement exam will tell you where to start reading.

Q1. What is the derivative of $\sin(x)$?

Q2. What is the second derivative of $A\sin(\omega x)$?

Q3. What is the value of x ?
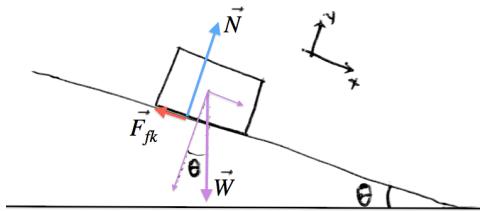
Q4. What is the magnitude of the gravitational force between two planets of mass M and mass m separated by a distance r?

Q5. Calculate $\lim\limits_{x\to 3^-}\dfrac{1}{x-3}$.

Q6. Solve for t in the equation:

$$7(3+4t) = 11(6t-4).$$

Q7. What is the component of the weight $\vec{W}$ acting in the x-direction?

Q8. A mass-spring system is undergoing simple harmonic motion. Its position function is $x(t) = A\sin(\omega t)$. What is its maximum acceleration?

---

[1] Ans: A1. $\cos(x)$, A2. $-A\omega^2\sin(\omega x)$, A3. $\frac{\sqrt{3}}{2}$, A4. $\|\vec{F_g}\| = \frac{GMm}{r^2}$, A5. $-\infty$, A6. $\frac{65}{38}$, A7. $+mg\sin\theta$, A8. $A\omega^2$. Key: If you didn't get Q3, Q6 right, you should read the book starting from Chapter 1. If you are mystified by Q1, Q2, Q5, read Chapter 5. If you want to learn how to solve Q4, Q7 and Q8, read Chapter 4.

Concept map

algebra — numbers

is the manipulation of

refer to

variables

are used in

equations

y=mx+b

quadratic eqn

inverse function

solve

systems of equations

algebra

y=ax²+bx+c | x² | √x

f(x-h)+k

|x|

exp(x) | ln(x)

sin(x) | asin(x)

cos(x) | acos(x)

tan(x) | atan(x)

function graphs

functions

triangle

trig identities

circle | radians | trigonometry

ellipse

parabola

hyperbola

geometry

high school math

have — components

vectors — are — directions

expressed w. respect to a

basis

=

coordinate system

motion

position — velocity — acceleration — F=ma — forces

tension | friction

gravity

spring

normal

linear motion

circular motion

angular motion

simple harmonic motion

potential energy

½mv²

kinetic energy — work

∫F·dx

mv

conservation — momentum

kinematics | **momentum** | **energy** | **dynamics**

vector operations

vector products — dot product

cross product

vectors

are similar to

complex numbers

mechanics

infinity

epsilon

delta

used in

limits

used in | used in | used in | used in

derivative operator

slope of the graph

derivative rules

higher derivatives

optimization

implicit differentiation

related by the

fundamental theorem

diff/int formulas

derivatives

integral operator

area under the graph

Riemann sum

techniques of integration

applications

integrals

sequences — sum — series

Taylor series

sequences and series

calculus

Figure 1: This diagram shows the connections between the concepts, topics, and subjects covered in the book. Seeing the connections between concepts is key to understanding math and physics. Consult the index on page 511 to find the exact location in the book where each concept is defined.

You can annotate the concept map with your current knowledge of each concept to keep track of your progress through the book.

- Add a single dot (•) next to all concepts you've heard of.
- Add two dots (••) next to concepts you think you know.
- Add three dots (•••) next to concepts you've used in exercises and problems.

By collecting some dots every week, you'll be able to move through the material in no time at all.

If you don't want to mark up your book, you can download a printable version of the concept map here: `bit.ly/mathphyscmap`.

Preface

This book contains lessons on topics in math and physics, written in a style that is jargon-free and to the point. Each lesson covers one concept at the depth required for a first-year university-level course. The main focus of this book is to highlight the intricate connections between the concepts of math and physics. Seeing the similarities and parallels between the concepts is the key to understanding.

Why?

The genesis of this book dates back to my student days when I was required to purchase expensive textbooks for my courses. Not only are these textbooks expensive, they are also tedious to read. Who has the energy to go through thousands of pages of explanations? I began to wonder, "What's the deal with these thick books?" Later, I realized mainstream textbooks are long because the textbook industry wants to make more profits. You don't need to read 1000 pages to learn calculus; the numerous full-page colour pictures and the repetitive text that are used to "pad" calculus textbooks are there to make the $200 price seem reasonable.

Looking at this situation, I said to myself, "Something must be done," and I sat down and wrote a modern textbook to explain math and physics concepts clearly, concisely, and affordably. There was no way I was going to let mainstream publishers ruin the learning experience of these beautiful subjects for the next generation of students.

How?

The sections in this book are **self-contained tutorials**. Each section covers the definitions, formulas, and explanations associated with a single topic. You can therefore read the sections in any order you find logical. Along the way, you will learn about the *connections* between the concepts of calculus and mechanics. Understanding mechanics is much easier if you know the ideas of calculus. At the same time, the

ideas behind calculus are best illustrated through concrete physics examples. Learning the two subjects simultaneously is the best approach.

In order to make the study of mechanics and calculus accessible for all readers, we'll begin with a review chapter on numbers, algebra, equations, functions, and other prerequisite concepts. If you feel a little rusty on those concepts, be sure to check out Chapter 1.

Each chapter ends with a section of **practice problems** designed to test your understanding of the concepts developed in that chapter. Make sure you spend plenty of time on these problems to practice what you've learned. Figuring out how to use an equation on your own in the process of solving a problem is a much more valuable experience than simply memorizing the equation.

For optimal learning efficiency, I recommend that you spend as much time working through the practice problems as you will spend reading the lessons. The problems you find difficult to solve will tell you which sections of the chapter you need to revisit. An additional benefit of testing your skills on the practice problems is that you'll be prepared in case a teacher ever tries to test you.

Throughout the book, I've included **links to internet resources** like animations, demonstrations, and webpages with further reading material. Once you understand the basics, you'll be able to understand far more internet resources. The links provided are a starting point for further exploration.

Is this book for you?

My aim is to make learning calculus and mechanics more accessible. Anyone should be able to open this book and become proficient in calculus and mechanics, regardless of their math background.

The book's primary intended audience is students. Students taking a mechanics class can read the chapters sequentially until Chapter 4, and optionally read Chapter 5 for bonus points. Taking a calculus course? Skip ahead directly to the calculus chapter (Chapter 5). High school students or university students taking a precalculus class will benefit from reading Chapter 1, which is a concise

MECH CLASS	CALC CLASS	PRECALC CLASS
Ch. 1	Ch. 1†	Ch. 1
Ch. 2	Ch. 2	Ch. 2†
Ch. 3		Ch. 3
Ch. 4		
Ch. 5†	Ch. 5	

$\dagger$ = optional reading.

but thorough review of fundamental math concepts like numbers, equations, functions, and trigonometry.

Non-students, don't worry: you don't need to be taking a class

in order to learn math. Independent learners interested in learning university-level material will find this book very useful. Many university graduates read this book to remember the calculus they learned back in their university days.

In general, anyone interested in rekindling their relationship with mathematics should consider this book as an opportunity to repair the broken connection. Math is good stuff; you shouldn't miss out on it. People who think they absolutely *hate* math should read Chapter 1 as therapy.

About the author

I have been tutoring math and physics for more than 17 years. Through this experience, I learned to break complicated ideas into smaller, interconnected chunks that are easy to understand. I think the best way to teach math and physics is to clearly define concepts and show the paths that connect them. It's not about how many equations you know, but about knowing how to get from one equation to another.

I completed my undergraduate studies at McGill University in electrical engineering, then did a M.Sc. in physics, and recently completed a Ph.D. in computer science. In my career as a researcher, I've been fortunate to learn from very inspirational teachers, who had the ability to distill the essential ideas and explain things in simple language. With my writing, I want to recreate the same learning experience for you. I founded Minireference Co. to revolutionize the textbook industry. We make textbooks that don't suck.

Ivan Savov
Montreal, 2020

Introduction

The last century has been marked by tremendous technological advances. Every sector of the economy has been transformed by the use of computers and the advent of the internet. There is no doubt technology's importance will continue to grow in the coming years.

The best part is that you don't need to know how technology works to use it. You need not understand how internet protocols operate to check your email and find original pirate material. You don't need to be a programmer to use computers to automate repetitive tasks and increase your productivity. However, when it comes to building *new* things, understanding how technology works becomes important. One particularly useful skill is the ability to **create mathematical models of real-world situations**. The techniques of mechanics and calculus are powerful building blocks for understanding the world around us. This is why these courses are taught in the first year of university studies: they contain keys that unlock the rest of science and engineering.

Calculus and mechanics can be difficult subjects. Understanding the material isn't hard *per se*, but it takes patience and practice to become comfortable with the new ideas. Calculus and mechanics become much easier to absorb when you break the material into manageable chunks. The concept map in Figure 1 (page v) shows an overview of all the concepts and topics we'll discuss in the book. There are a lot of new things to learn, but don't worry—we'll navigate the material step by step and it will all make sense in the end.

Before we start with the equations, it's worthwhile to preview the material covered in this book. After all, you should know what kind of trouble you're getting yourself into.

Chapter 1 is a comprehensive review of math fundamentals including algebra, equations, functions, geometry, and trigonometry. The exposition of each topic is brief to make for easy reading. This chapter is highly recommended for readers who haven't looked at math recently; if you need a refresher on math, Chapter 1 is for you. It is extremely important to firmly grasp the basics. What is $\sin(0)$?

What is $\sin(\pi/4)$? What does the graph of $\sin(x)$ look like? Adult learners can use this review chapter as therapy to recover from any traumatizing math learning experiences they may have encountered in high school.

In Chapter 2, we'll look at how techniques of high school math can be used to describe and model the world. We'll learn about the basic laws that govern the motion of objects and the mathematical equations that describe the motion. By the end of this chapter, you'll understand the concepts of velocity and acceleration, and be able to predict the flight time of a ball thrown in the air.

In Chapter 3, we'll learn about vectors. Vectors describe directional quantities like forces and velocities. We need vectors to properly understand the laws of physics. Vectors are used in many areas of science and technology, so becoming comfortable with vector calculations will pay dividends when learning other subjects.

Chapter 4 is all about mechanics. We'll study the motion of objects, predict their future trajectories, and learn how to use abstract concepts like momentum and energy. Science students who "hate" physics can study this chapter to learn how to use the 20 main equations and laws of physics. You'll see physics is actually quite simple.

Chapter 5 covers topics from differential calculus and integral calculus. We'll study limits, derivatives, integrals, sequences, and series. You'll find that 130 pages are enough to cover all the concepts in calculus, including practical applications.
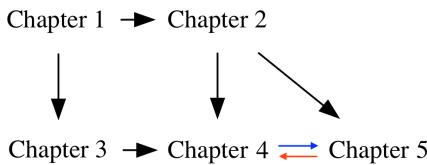
Figure 2: The prerequisite structure for the chapters in this book.

Calculus and mechanics are often taught as separate subjects. It shouldn't be like that! If you learn calculus without mechanics, it will be boring. If you learn physics without calculus, you won't truly understand. The exposition in this book covers both subjects in an integrated manner and highlights the connections between them.

Are you ready for this? Let's dig in!

Chapter 1

Math fundamentals

In this chapter we'll review the fundamental ideas of mathematics, including numbers, equations, and functions. To understand college-level textbooks, you need to be comfortable with mathematical calculations. Many people have trouble with math, however. Some people say they *hate* math, or could never learn it. It's not uncommon for children who score poorly on their school math exams to develop math complexes in their grown lives. If you are carrying any such emotional baggage, you can drop it right here and right now.

Do NOT worry about math! You are an adult, and you can learn math much more easily than when you were a kid. We'll review *everything* you need to know about high school math, and by the end of this chapter, you'll see that math is nothing to worry about.
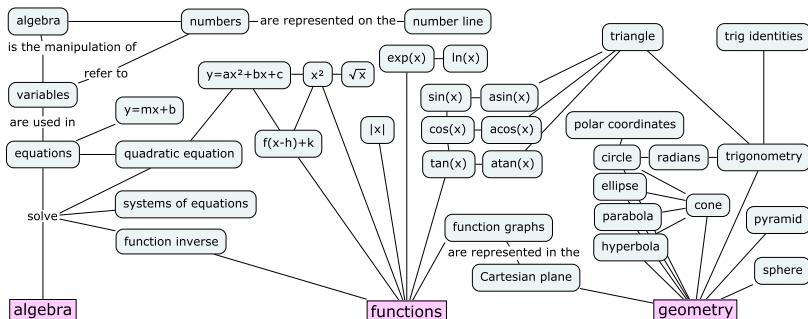

Figure 1.1: A concept map showing the mathematical topics that we will cover in this chapter. We'll learn how to solve equations using algebra, how to model the world using functions, and how to think geometrically. The material in this chapter is required for your understanding of the more advanced topics in this book.

1.1 Solving equations

Most math skills boil down to being able to manipulate and solve equations. Solving an equation means finding the value of the unknown in the equation.

Check this shit out:

$$x^2 - 4 = 45.$$

To solve the above equation is to answer the question "What is x?" More precisely, we want to find the number that can take the place of x in the equation so that the equality holds. In other words, we're asking,

"Which number times itself minus four gives 45?"

That is quite a mouthful, don't you think? To remedy this verbosity, mathematicians often use specialized symbols to describe math operations. The problem is that these specialized symbols can be very confusing. Sometimes even the simplest math concepts are inaccessible if you don't know what the symbols mean.

What are your feelings about math, dear reader? Are you afraid of it? Do you have anxiety attacks because you think it will be too difficult for you? Chill! Relax, my brothers and sisters. There's nothing to it. Nobody can magically guess the solution to an equation immediately. To find the solution, you must break the problem into simpler steps. Let's walk through this one together.

To find x, we can manipulate the original equation, transforming it into a different equation (as true as the first) that looks like this:

$$x = \text{only numbers.}$$

That's what it means to *solve* an equation: the equation is solved because the unknown is isolated on one side, while the constants are grouped on the other side. You can type the numbers on the right-hand side into a calculator and obtain the numerical value of x.

By the way, before we continue our discussion, let it be noted: the equality symbol ($=$) means that all that is to the left of $=$ is equal to all that is to the right of $=$. To keep this equality statement true, **for every change you apply to the left side of the equation, you must apply the same change to the right side of the equation**.

To find x, we need to manipulate the original equation into its final form, simplifying it step by step until it can't be simplified any further. The only requirement is that the manipulations we make transform one true equation into another true equation. In this example, the first simplifying step is to add the number four to both sides of the equation:

$$x^2 - 4 \; + 4 = 45 \; + 4,$$

which simplifies to
$$x^2 = 49.$$

Now the expression looks simpler, yes? How did I know to perform this operation? I wanted to "undo" the effects of the operation -4. We undo an operation by applying its *inverse*. In the case where the operation is the subtraction of some amount, the inverse operation is the addition of the same amount. We'll learn more about function inverses in Section 1.5.

We're getting closer to our goal of *isolating* x on one side of the equation, leaving only numbers on the other side. The next step is to undo the square x^2 operation. The inverse operation of squaring a number x^2 is to take its square root $\sqrt{\ }$, so that's what we'll do next. We obtain
$$\sqrt{x^2} = \sqrt{49}.$$

Notice how we applied the square root to both sides of the equation? If we don't apply the same operation to both sides, we'll break the equality!

The equation $\sqrt{x^2} = \sqrt{49}$ simplifies to
$$|x| = 7.$$

What's up with the vertical bars around x? The notation $|x|$ stands for the *absolute value* of x, which is the same as x except we ignore the sign that indicates whether x is positive or negative. For example $|5| = 5$ and $|-5| = 5$, too. The equation $|x| = 7$ indicates that both $x = 7$ and $x = -7$ satisfy the equation $x^2 = 49$. Seven squared is 49, $7^2 = 49$, and negative seven squared is also 49, $(-7)^2 = 49$, because the two negative signs cancel each other out.

The final solutions to the equation $x^2 - 4 = 45$ are
$$x = 7 \quad \text{and} \quad x = -7.$$

Yes, there are *two* possible answers. You can check that both of the above values of x satisfy the initial equation $x^2 - 4 = 45$.

If you are comfortable with all the notions of high school math and you feel you could have solved the equation $x^2 - 4 = 45$ on your own, then you can skim through this chapter quickly. If on the other hand you are wondering how the squiggle killed the power two, then this chapter is for you! In the following sections we will review all the essential concepts from high school math that you will need to power through the rest of this book. First, let me tell you about the different kinds of numbers.

1.2 Numbers

In the beginning, we must define the main players in the world of math: numbers.

Definitions

Numbers are the basic objects we use to count, measure, quantify, and calculate things. Mathematicians like to classify the different kinds of number-like objects into categories called *sets*:

- The natural numbers: $\mathbb{N} = \{0, 1, 2, 3, 4, 5, 6, 7, \ldots\}$
- The integers: $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$
- The rational numbers: $\mathbb{Q} = \{\frac{5}{3}, \frac{22}{7}, 1.5, 0.125, -7, \ldots\}$
- The real numbers: $\mathbb{R} = \{-1, 0, 1, \sqrt{2}, e, \pi, 4.94\ldots, \ldots\}$
- The complex numbers: $\mathbb{C} = \{-1, 0, 1, i, 1+i, 2+3i, \ldots\}$

These categories of numbers should be somewhat familiar to you. Think of them as neat classification labels for everything that you would normally call a number. Each group in the above list is a *set*. A set is a collection of items of the same kind. Each collection has a name and a precise definition for which items belong in that collection. Note also that each of the sets in the list contains all the sets above it, as illustrated in Figure 1.2. For now, we don't need to go into the details of sets and set notation, but we do need to be aware of the different sets of numbers.
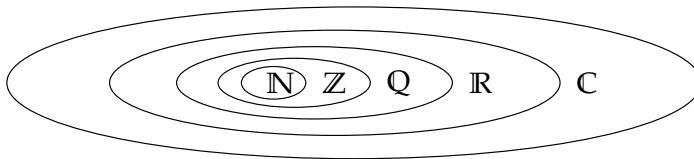
Figure 1.2: An illustration of the nested containment structure of the different number sets. The set of natural numbers is contained in the set of integers, which in turn is contained in the set of rational numbers. The set of rational numbers is contained in the set of real numbers, which is contained in the set of complex numbers.

Why do we need so many different sets of numbers? Each set of numbers is associated with more and more advanced mathematical problems.

The simplest numbers are the natural numbers $\mathbb{N}$, which are sufficient for all your math needs if all you're going to do is *count* things. How many goats? Five goats here and six goats there so the total is

number 2 corresponds to a line segment of length two, which is a geometric object in a one-dimensional space. If we add a line segment of length two in a second dimension, we obtain a square with area 2^2 in a two-dimensional space. Adding a third dimension, we obtain a cube with volume 2^3 in a three-dimensional space. Indeed, raising a base a to the exponent 2 is commonly called "a squared," and raising a to the power of 3 is called "a cubed."

The geometrical analogy about one-dimensional quantities as lengths, two-dimensional quantities as areas, and three-dimensional quantities as volumes is good to keep in mind.

$$2^1 = 2 \qquad 2^2 = 4 \qquad 2^3 = 8$$
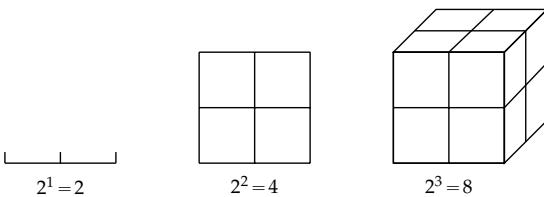
Figure 1.5: Geometric interpretation for exponents 1, 2, and 3. A length raised to exponent 2 corresponds to the area of a square. The same length raised to exponent 3 corresponds to the volume of a cube.

Our visual intuition works very well up to three dimensions, but we can use other means of visualizing higher exponents, as demonstrated in Figure 1.6.

Operator precedence

There is a standard convention for the order in which mathematical operations must be performed. The basic algebra operations have the following precedence:

1. Parentheses
2. Exponents
3. Multiplication and Division
4. Addition and Subtraction

If you're seeing this list for the first time, the acronym PEMDAS and the associated mnemonic "Please Excuse My Dear Aunt Sally," might help you remember the order of operations.

For instance, the expression $5 \cdot 3^2 + 13$ is interpreted as "First find the square of 3, then multiply it by 5, and then add 13." Parentheses are needed to carry out the operations in a different order: to multiply 5 times 3 first and *then* take the square, the equation should read
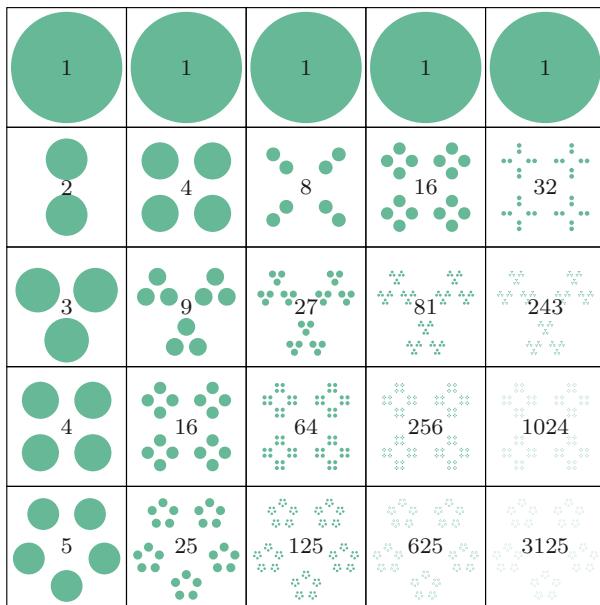
Figure 1.6: Visualization of numbers raised to different exponents. Each box in this grid contains a^n dots, where the base a varies from one through five, and the exponent n varies from one through five. In the first row we see that the number $a = 1$ raised to any exponent is equal to itself. The second row corresponds to the base $a = 2$ so the number of dots doubles each time we increase the exponent by one. Starting from $2^1 = 2$ in the first column, we end up with $2^5 = 32$ in the last column. The rest of the rows show how exponentiation works for different bases.

$(5 \cdot 3)^2 + 13$, where parentheses indicate that the square acts on $(5 \cdot 3)$ as a whole and not on 3 alone.

Exercises

E1.1 Solve for the unknown x in the following equations:

    a) $3x + 2 - 5 = 4 + 2$ **b)** $\frac{1}{2}x - 3 = \sqrt{3} + 12 - \sqrt{3}$

    c) $\frac{7x-4}{2} + 1 = 8 - 2$ **d)** $5x - 2 + 3 = 3x - 5$

E1.2 Indicate all the number sets the following numbers belong to.

    a) -2 **b)** $\sqrt{-3}$ **c)** $8 \div 4$ **d)** $\frac{5}{3}$ **e)** $\frac{\pi}{2}$

E1.3 Calculate the values of the following expressions:

    a) $2^3 3 - 3$ **b)** $2^3(3 - 3)$ **c)** $\frac{4-2}{3^3}(6 \cdot 7 - 41)$
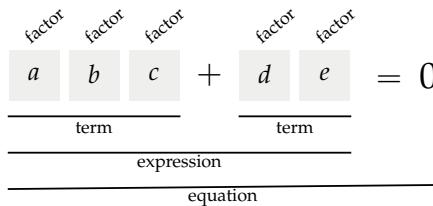
Figure 1.14: Diagram showing the names used to describe the different parts of the equation $abc + de = 0$.

We use the distributive property every time we *expand* brackets. For example $a(b + c + d) = ab + ac + ad$. The brackets, also known as parentheses, indicate the expression $(b + c + d)$ must be treated as a whole; as a factor consisting of three terms. Multiplying this expression by a is the same as multiplying each term by a.

The opposite operation of expanding is called *factoring*, which consists of rewriting the expression with the common parts taken out in front of a bracket: $ab + ac = a(b + c)$. In this section, we'll discuss all algebra operations and illustrate what they're capable of.

Example   Suppose we are asked to solve for t in the equation

$$7(3 + 4t) = 11(6t - 4).$$

Since the unknown t appears on both sides of the equation, it is not immediately obvious how to proceed.

To solve for t, we can bring all t terms to one side and all constant terms to the other side. First, expand the two brackets to obtain

$$21 + 28t = 66t - 44.$$

Then move things around to relocate all ts to the equation's right-hand side and all constants to the left-hand side:

$$21 + 44 = 66t - 28t.$$

We see t is contained in both terms on the right-hand side, so we can "factor it out" by rewriting the equation as

$$21 + 44 = t(66 - 28).$$

The answer is within close reach: $t = \frac{21+44}{66-28} = \frac{65}{38}$.

Expanding brackets

To *expand* a bracket is to multiply each term inside the bracket by the factor outside the bracket. The key thing to remember when expanding brackets is to apply the *distributive* property: $a(x + y) = ax + ay$.

constant correction factor $(x + p)^2 + k$. This algebra technique was described in one of the first books on *al-jabr* (algebra), written by Al-Khwarizmi around the year 800 CE. The name "completing the square" comes from the ingenious geometric construction used by this procedure. Yes, we can use geometry to solve algebra problems!

We assume the starting point for the procedure is a quadratic expression whose quadratic coefficient is one, $1x^2 + Bx + C$, and use capital letters B and C to denote the linear and constant coefficients. The capital letters are to avoid any confusion with the quadratic expression $ax^2 + bx + c$, for which $a \neq 1$. Note we can always write $ax^2 + bx + c$ as $a(x^2 + \frac{b}{a}x + \frac{c}{a})$ and apply the procedure to the expression inside the brackets, identifying $\frac{b}{a}$ with B and $\frac{c}{a}$ with C.

First let's rewrite the quadratic expression $x^2 + Bx + C$ by splitting the linear term into two equal parts:

$$x^2 + \tfrac{B}{2}x + \tfrac{B}{2}x + C.$$

We can interpret the first three terms geometrically as follows: the x^2 term corresponds to a square with side length x, while the two $\frac{B}{2}x$ terms correspond to rectangles with sides $\frac{B}{2}$ and x. See the left side of Figure 1.15 for an illustration.
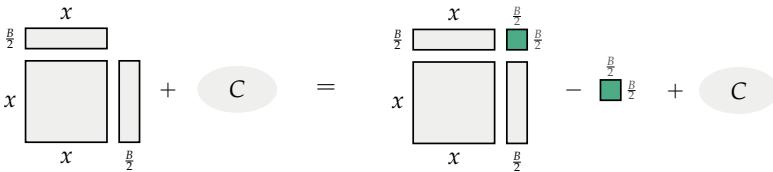
Figure 1.15: To complete the square in the expression $x^2 + Bx + C$, we need to add the quantity $(\frac{B}{2})^2$, which corresponds to a square (shown in darker colour) with sides equal to half the coefficient of the linear term. We also subtract $(\frac{B}{2})^2$ so the overall value of the expression remains unchanged.

The square with area x^2 and the two rectangles can be positioned to form a larger square with side length $(x + \frac{B}{2})$. Note there's a small piece of sides $\frac{B}{2}$ by $\frac{B}{2}$ missing from the corner. To *complete the square*, we can add a term $(\frac{B}{2})^2$ to this expression. To preserve the equality, we also subtract $(\frac{B}{2})^2$ from the expression to obtain:

$$
\begin{aligned}
x^2 + \tfrac{B}{2}x + \tfrac{B}{2}x + C &= \underbrace{x^2 + \tfrac{B}{2}x + \tfrac{B}{2}x + (\tfrac{B}{2})^2}_{} - (\tfrac{B}{2})^2 + C \\
&= (x + \tfrac{B}{2})^2 - (\tfrac{B}{2})^2 + C.
\end{aligned}
$$

The right-hand side of this equation describes the area of the square with side length $(x + \frac{B}{2})$, minus the area of the small square $(\frac{B}{2})^2$, plus the constant C, as illustrated on the right side of Figure 1.15.

Exercises

E1.7 Factor the following quadratic expressions:

    a) $x^2 - 8x + 7$ **b)** $x^2 + 4x + 4$ **c)** $x^2 - 9$

Hint: Guess the values p and q in the expression $(x + p)(x + q)$.

E1.8 Solve the equations by completing the square.

    a) $x^2 + 2x - 15 = 0$ **b)** $x^2 + 4x + 1 = 0$

1.7   Solving quadratic equations

What would you do if asked to solve for x in the quadratic equation $2x^2 = 4x + 6$? This is called a *quadratic equation* since it contains the unknown variable x squared. The name comes from the Latin *quadratus*, which means square. Quadratic equations appear often, so mathematicians created a general formula for solving them. In this section, we'll learn about this formula and use it to put some quadratic equations in their place.

Before we can apply the formula, we need to rewrite the equation we are trying to solve in the following form:

$$ax^2 + bx + c = 0.$$

This is called the *standard form* of the quadratic equation. We obtain this form by moving all the numbers and xs to one side and leaving only 0 on the other side. For example, to transform the quadratic equation $2x^2 = 4x + 6$ into standard form, we subtract $4x + 6$ from both sides of the equation to obtain $2x^2 - 4x - 6 = 0$. What are the values of x that satisfy this equation?

Quadratic formula

> The solutions to the equation $ax^2 + bx + c = 0$ for $a \neq 0$ are
>
> $$x_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \quad \text{and} \quad x_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}.$$

The quadratic formula is usually abbreviated $x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$, where the sign "$\pm$" stands for both "$+$" and "$-$." The notation "$\pm$" allows us to express both solutions x_1 and x_2 in one equation, but you should keep in mind there are really two solutions.

1.9 Logarithms

Some people think the word "logarithm" refers to some mythical, mathematical beast. Legend has it that logarithms are many-headed, breathe fire, and are extremely difficult to understand. Nonsense! Logarithms are simple. It will take you at most a couple of pages to get used to manipulating them, and that is a good thing because logarithms are used all over the place.

The strength of your sound system is measured in logarithmic units called decibels. This is because your ears are sensitive only to exponential differences in sound intensity. Logarithms allow us to compare very large numbers and very small numbers on the same scale. If sound were measured in linear units instead of logarithmic units, your sound system's volume control would need to range from 1 to $1\,048\,576$. That would be weird, no? This is why we use the logarithmic scale for volume notches. Using a logarithmic scale, we can go from sound intensity level 1 to sound intensity level $1\,048\,576$ in 20 "progressive" steps. Assume each notch doubles the sound intensity, rather than increasing the intensity by a fixed amount. If the first notch corresponds to 2, the second notch is 4—still probably inaudible, turn it up! By the time you get to the sixth notch you're at $2^6 = 64$ sound intensity, which is the level of audible music. The tenth notch corresponds to sound intensity $2^{10} = 1024$ (medium-strength sound), and finally the twentieth notch reaches a max power of $2^{20} = 1\,048\,576$, at which point the neighbours come to complain.

Definitions

You're hopefully familiar with these following concepts from the previous section:

- b^x: the exponential function base b
- $\exp(x) = e^x$: the exponential function base e, Euler's number
- 2^x: exponential function base 2
- f: the notion of a function
- f^{-1}: the inverse function of f. The inverse function is defined in terms of f such that $f^{-1}(f(x)) = x$. In other words, if you apply f to some number x and get the output y, and then you pass y through f^{-1}, the output will be x again. The inverse function f^{-1} undoes the effects of the function f.

In this section we'll play with the following new concepts:

- $\log_b(x)$: the logarithm of x base b is the inverse function of b^x.
- $\ln(x)$: the "natural" logarithm base e. This is the inverse of e^x.
- $\log_2(x)$: the logarithm base 2 is the inverse of 2^x.

variable (the one that varies freely), and the y-axis represents the *dependent* variable $f(x)$, since $f(x)$ depends on x.

To draw the graph of any function $f(x)$, use the following procedure. Imagine making a sweep over all of the possible input values for the function. For each input x, put a point at the coordinates $(x, y) = (x, f(x))$ in the Cartesian plane. Using the graph of a function, you can literally *see* what the function does: the "height" y of the graph at a given x-coordinate tells you the value of the function $f(x)$.

Dimensions

The number line is one-dimensional. Every number x can be visualized as a point on the number line. The Cartesian plane has two dimensions: the x dimension and the y dimension. If we need to visualize math concepts in 3D, we can use a three-dimensional coordinate system with x, y, and z axes (see Figure 3.10 on page 196).

1.11   Functions

We need to have a relationship talk. We need to talk about functions. We use functions to describe the relationships between variables. In particular, functions describe how one variable *depends* on another.

For example, the revenue R from a music concert depends on the number of tickets sold n. If each ticket costs \$25, the revenue from the concert can be written *as a function of n* as follows: $R(n) = 25n$. Solving for n in the equation $R(n) = 7000$ tells us the number of ticket sales needed to generate \$7000 in revenue. This is a simple model of a function; as your knowledge of functions builds, you'll learn how to build more detailed models of reality. For instance, if you need to include a 5% processing charge for issuing the tickets, you can update the revenue model to $R(n) = 0.95 \cdot 25 \cdot n$. If the estimated cost of hosting the concert is $C = \$2000$, then the profit from the concert P can be modelled as

$$P(n) = R(n) - C$$
$$= 0.95 \cdot \$25 \cdot n - \$2000$$

The function $P(n) = 23.75n - 2000$ models the profit from the concert as a function of the number of tickets sold. This is a pretty good model already, and you can always update it later as you learn more information.

The more functions you know, the more tools you have for modelling reality. To "know" a function, you must be able to understand

and connect several of its aspects. First you need to know the function's mathematical **definition**, which describes exactly what the function does. Starting from the function's definition, you can use your existing math skills to find the function's **properties**. You must also know the **graph** of the function; what the function looks like if you plot x versus $f(x)$ in the Cartesian plane. It's also a good idea to remember the **values** of the function for some important inputs. Finally—and this is the part that takes time—you must learn about the function's **relations** to other functions.

Definitions

A *function* is a mathematical object that takes numbers as inputs and produces numbers as outputs. We use the notation

$$f\colon A \to B$$

to denote a function from the input set A to the output set B. In this book, we mostly study functions that take real numbers as inputs and give real numbers as outputs: $f\colon \mathbb{R} \to \mathbb{R}$.
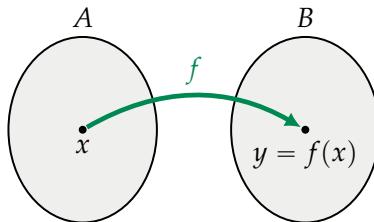
Figure 1.20: An abstract representation of a function f from the set A to the set B. The function f is the arrow which *maps* each input x in A to an output $f(x)$ in B. The output of the function $f(x)$ is also denoted y.

A function is not a number; rather, it is a *mapping* from numbers to numbers. We say "f maps x to $f(x)$." For any input x, the output value of f for that input is denoted $f(x)$, which is read as "f of x."

We'll now define some fancy technical terms used to describe the input and output sets of functions.

- A: the *source set* of the function describes the types of numbers that the function takes as inputs.
- $\mathrm{Dom}(f)$: the *domain* of a function is the set of allowed input values for the function.

To build mathematical intuition, it's essential you understand functions' graphs. Memorizing the definitions and properties of functions gets a lot easier with visual accompaniment. Indeed, remembering what the function "looks like" is a great way to train yourself to recognize various types of functions. Figure 1.28 shows the graphs of some of the most important functions we'll use in this book.
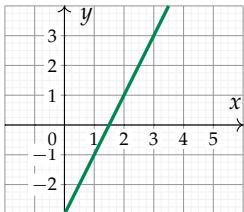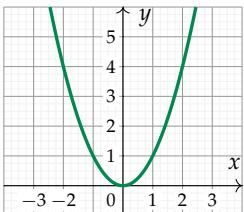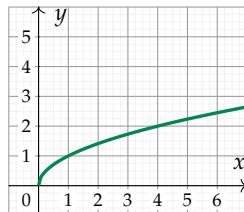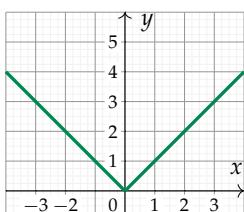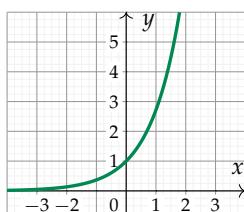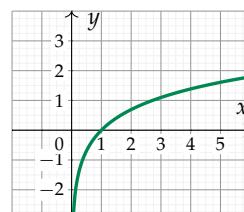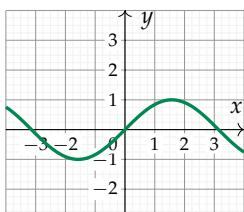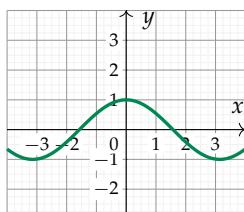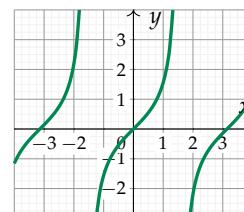
(a) $f(x) = 2x - 3$ (b) $f(x) = x^2$ (c) $f(x) = \sqrt{x}$

(d) $f(x) = |x|$ (e) $f(x) = e^x$ (f) $f(x) = \ln(x)$

(g) $f(x) = \sin(x)$ (h) $f(x) = \cos(x)$ (i) $f(x) = \tan(x)$

Figure 1.28: We'll see many types of function graphs in the next pages.

Line

The equation of a line describes an input-output relationship where the change in the output is *proportional* to the change in the input. The equation of a line is

$$f(x) = mx + b.$$

Square root

The square root function is denoted

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

The square root $\sqrt{x}$ is the inverse function of the square function x^2 when the two functions are defined as $f : \mathbb{R}_+ \to \mathbb{R}_+$. The symbol $\sqrt{c}$ refers to the *positive* solution of $x^2 = c$. Note that $-\sqrt{c}$ is also a solution of $x^2 = c$.

Graph
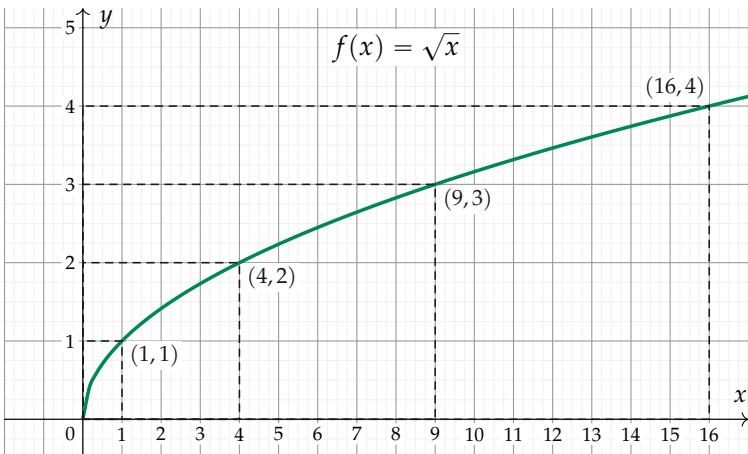
Figure 1.31: The graph of the function $f(x) = \sqrt{x}$. The domain of the function is $\mathbb{R}_+$ because we can't take the square root of a negative number.

Properties

- Domain: $\mathbb{R}_+ = \{x \in \mathbb{R} \mid x \geqslant 0\}$. The function $f(x) = \sqrt{x}$ is only defined for nonnegative inputs. There is no real number y such that y^2 is negative, hence the function $f(x) = \sqrt{x}$ is not defined for negative inputs x.
- Image: $\mathbb{R}_+ = \{y \in \mathbb{R} \mid y \geqslant 0\}$. The outputs of the function $f(x) = \sqrt{x}$ are nonnegative numbers since $\sqrt{x} \geqslant 0$.

In addition to *square* root, there is also *cube* root $f(x) = \sqrt[3]{x} = x^{\frac{1}{3}}$, which is the inverse function for the cubic function $f(x) = x^3$. We have $\sqrt[3]{8} = 2$ since $2 \times 2 \times 2 = 8$. More generally, we can define the nth-root function $\sqrt[n]{x}$ as the inverse function of x^n.

Natural logarithm

The natural logarithm function is denoted

$$f(x) = \ln(x) = \log_e(x).$$

The function $\ln(x)$ is the inverse function of the exponential e^x.

Graph
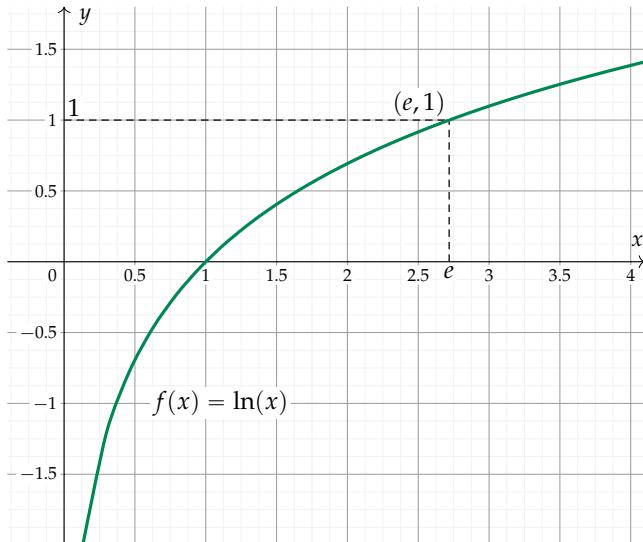
Figure 1.39: The graph of the function $\ln(x)$ passes through the following coordinates: $(\frac{1}{e^2}, -2)$, $(\frac{1}{e}, -1)$, $(1, 0)$, $(e, 1)$, $(e^2, 2)$, $(e^3, 3)$, $(e^4, 4)$, etc.

Properties

- Domain: $\{x \in \mathbb{R} \mid x > 0\}$
- Image: $\mathbb{R}$

Exercises

E1.17 Find the domain, the image, and the roots of $f(x) = 2\cos(x)$.

E1.18 What are the degrees of the following polynomials? Are they even, odd, or neither?

    a) $p(x) = x^2 - 5x^4 + 1$ **b)** $q(x) = x - x^3 + x^5 - x^7$

E1.19 Solve for x in the following polynomial equations.

    a) $3x + x^2 = x - 15 + 2x^2$ **b)** $3x^2 - 4x - 4 + x^3 = x^3 + 2x + 2$

1.15 Trigonometry

If one of the angles in a triangle is equal to $90°$, we call this triangle a *right-angle triangle*. In this section we'll discuss right-angle triangles in great detail and get to know their properties. We'll learn some fancy new terms like *hypotenuse*, *opposite*, and *adjacent*, which are used to refer to the different sides of a triangle. We'll also use the functions *sine*, *cosine*, and *tangent* to compute the *ratios of lengths* in right triangles.

Understanding triangles and their associated trigonometric functions is of fundamental importance: you'll need this knowledge for your future understanding of mathematical concepts like vectors and complex numbers.
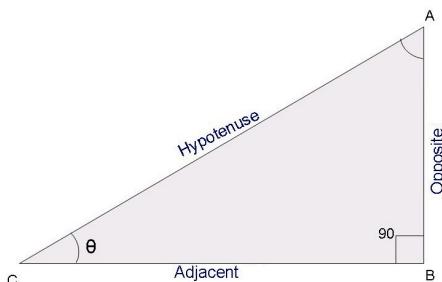
Figure 1.54: A right-angle triangle. The angle at the base is denoted θ and the names of the sides of the triangle are indicated.

Concepts

- A, B, C: the three *vertices* of the triangle
- θ: the angle at the vertex C. Angles can be measured in degrees or radians.
- opp $= AB$: the length of the *opposite* side to θ
- adj $= BC$: the length of side *adjacent* to θ
- hyp $= AC$: the *hypotenuse*. This is the triangle's longest side.
- h: the "height" of the triangle (in this case $h = $ opp $= AB$)
- $\sin\theta = \frac{\text{opp}}{\text{hyp}}$: the *sine* of theta is the ratio of the length of the opposite side and the length of the hypotenuse
- $\cos\theta = \frac{\text{adj}}{\text{hyp}}$: the *cosine* of theta is the ratio of the adjacent length and the hypotenuse length
- $\tan\theta = \frac{\sin\theta}{\cos\theta} = \frac{\text{opp}}{\text{adj}}$: the *tangent* is the ratio of the opposite length divided by the adjacent length
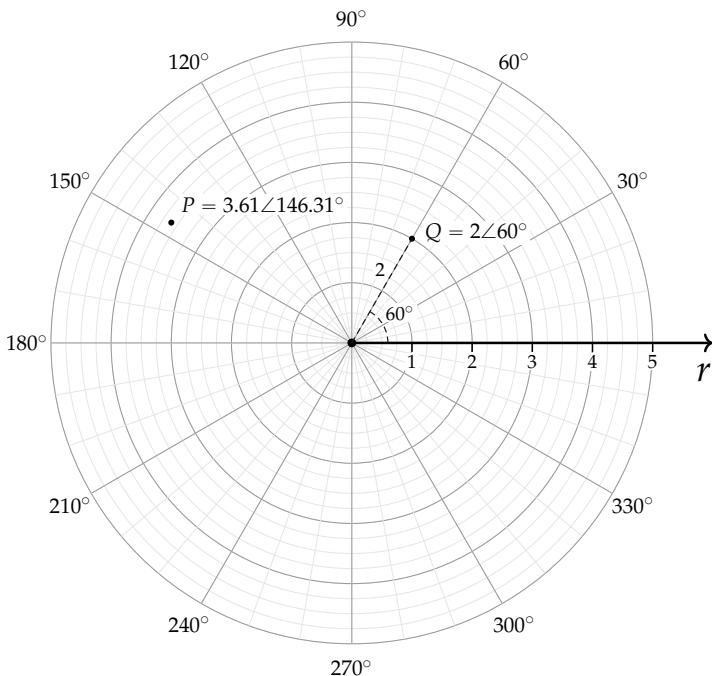
Figure 1.61: We can use the polar coordinate system to describe points in the two-dimensional plane. The polar coordinates $r\angle\theta$ describe the point located at the distance r from the origin in the direction θ.

The point $Q = 2\angle 60°$ is equally described by the polar coordinates $2\angle - 300°$, since a clockwise turn of $300°$ is the same as a counter-clockwise turn of $60°$. We can also describe the same point Q using the polar coordinates $-2\angle 240°$ and $-2\angle - 120°$, which tell us to turn in the direction opposite to $60°$ and measure a negative distance $r = -2$. While all of these polar coordinates for Q are equivalent, the preferred way to specify polar coordinates is with positive r values and angles $|\theta| \leqslant 180°$.

Converting between Cartesian and polar coordinates

Figure 1.62 shows a point whose location is described both in terms of Cartesian coordinates (x, y) and polar coordinates $r\angle\theta$. The triangle formed by the coordinates $(0,0)$, $(x,0)$, and (x,y) is a right-angle triangle. This means we can apply our knowledge of the trigonometric functions sin, cos, and tan to obtain formulas for converting between Cartesian coordinates (x, y) and polar coordinates $r\angle\theta$.

To convert from polar coordinates $r\angle\theta$ to (x, y) coordinates, we

a circle. When $\varepsilon = 0$ the ellipse is a circle with radius a, and both focal points are located at the centre. As the eccentricity ε increases, the ellipse becomes more elongated and the focal points spread farther apart.

Polar coordinates

Consider a polar coordination system whose centre is located at the focus F_2. We can describe the ellipse by specifying the function $r_2(\theta)$, which describes the distance from the focus F_2 to the point E on the ellipse as a function of the angle θ (see Figure 1.67). Recall that for functions in polar coordinates, the angle θ is the independent variable that varies from 0 to 2π (360°), and the dependent variable is the distance $r_2(\theta)$.
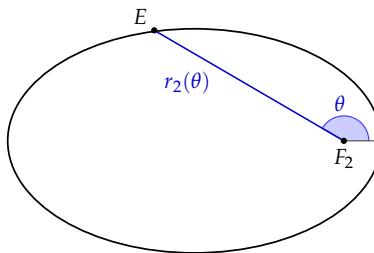
Figure 1.67: The function $r_2(\theta)$ in polar coordinates specifies the distance between the point E on the ellipse and the focal point F_2 for all angles.

The function that describes an ellipse in polar coordinates is

$$r_2(\theta) = \frac{a(1 - \varepsilon^2)}{1 + \varepsilon \cos(\theta)},$$

where the angle θ is measured with respect to the semi-major axis. The distance is smallest when $\theta = 0$ with $r_2(0) = a - c = a(1 - \varepsilon)$ and largest when $\theta = \pi$ with $r_2(\pi) = a + c = a(1 + \varepsilon)$.

Calculating the orbit of the Earth

The motion of the Earth around the Sun is an ellipse with the Sun positioned at the focus F_2. We can therefore use the polar coordinates formula $r_2(\theta)$ to describe the distance of the Earth from the Sun. The eccentricity of Earth's orbit around the Sun is $\varepsilon = 0.01671123$, and the half-length of the major axis is $a = 149\,598\,261$ km. We substitute these values into the general formula for $r_2(\theta)$ and obtain the following equation:

$$r_2(\theta) = \frac{149\,556\,484}{1 + 0.01671123 \cos(\theta)} \quad \text{km.}$$

The point where the Earth is closest to the Sun is called the *perihelion*. It occurs when $\theta = 0$, which happens around the 3rd of January. The moment where the Earth is most distant from the Sun is called the *aphelion* and corresponds to the angle $\theta = \pi$. Earth's *aphelion* happens around the 3rd of July.

Let's use the formula for $r_2(\theta)$ to predict the *perihelion* and *aphelion* distances of Earth's orbit:

$$r_{2,\text{peri}} = r_2(0) = \frac{149556483}{1 + 0.01671123 \cos(0)} = 147\,098\,290 \text{ km,}$$

$$r_{2,\text{aphe}} = r_2(\pi) = \frac{149556483}{1 + 0.01671123 \cos(\pi)} = 152\,098\,232 \text{ km.}$$

Google "perihelion" and "aphelion" to verify that the above predictions are accurate. It's kind of cool that a mathematical formula can describe the motion of our planet, don't you think?
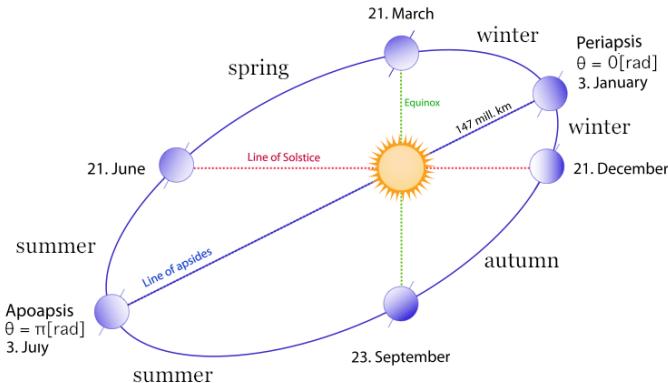
Figure 1.68: The orbit of the Earth around the Sun. Key points of the orbit are labelled. The seasons in the Northern hemisphere are also indicated.

The angle θ of the Earth relative to the Sun can be described as a function of time $\theta(t)$. The exact formula of the function $\theta(t)$ that describes the angle as a function of time is fairly complicated, so we won't go into the details. Let's simply look at the values of $\theta(t)$ with t measured in days shown in Table 1.3. We'll begin on Jan 3rd.

Newton's insight

Contrary to common belief, Newton did not discover his theory of gravitation because an apple fell on his head while sitting under a tree. What actually happened is that he started from Kepler's laws

Don't worry about $\cosh\mu$ and $\sinh\mu$ too much. The hyperbolic trig functions are used much less often than the circular trigonometric functions $\cos\theta$ and $\sin\theta$. The main thing to remember is the general pattern: cosine functions are used to denote horizontal coordinates and sine functions are used to denote vertical coordinates.

The conic sections

There is a deep connection between the geometric shapes of the circle, the ellipse, the parabola, and the hyperbola. These seemingly different shapes can be obtained, geometrically speaking, from a single object: the cone. We can obtain the four curves by slicing the cone at different angles, as illustrated in Figure 1.75.

circle   ellipse   parabola   hyperbola
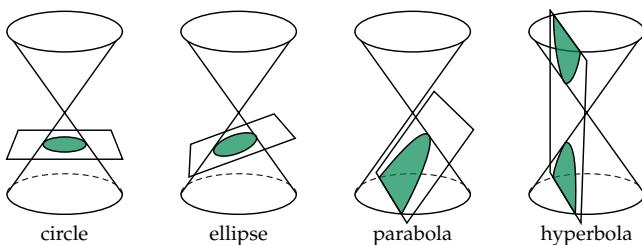
Figure 1.75: Taking slices through a cone at different angles produces different geometric shapes: a circle, an ellipse, a parabola, or a hyperbola.

Conic sections in polar coordinates

All four conic sections can be described by the same function in polar coordinates:

$$r(\theta) = \frac{q(1+\varepsilon)}{1+\varepsilon\cos(\theta)},$$

where q is the curve's closest distance to a focal point and ε is the curve's eccentricity. For a circle, $q = R$ (the radius) and the eccentricity parameter is $\varepsilon = 0$. For an ellipse, $q = a(1-\varepsilon)$ and the eccentricity parameter varies between 0 and 1 ($0 \leqslant \varepsilon < 1$). Note we include the case $\varepsilon = 0$ since a circle is a special case of an ellipse. For a parabola, $q = f$ (the focal length) and the eccentricity is $\varepsilon = 1$. For a hyperbola, $q = a(\varepsilon - 1)$ and the eccentricity is $\varepsilon > 1$.

We can use the eccentricity parameter ε to classify all four curves. Depending on the value of ε, the equation $r(\theta)$ defines either a circle, an ellipse, a parabola, or a hyperbola. Table 1.5 summarizes all our observations regarding conic sections.

The motion of the planets is explained by Newton's law of gravitation. The gravitational interaction between two bodies always

Conic section	Equation	Polar function	Eccentricity
Circle	$x^2 + y^2 = R^2$	$r(\theta) = R$	$\varepsilon = 0$
Ellipse	$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$	$r(\theta) = \frac{a(1-\varepsilon^2)}{1+\varepsilon\cos(\theta)}$	$\varepsilon = \sqrt{1 - \frac{b^2}{a^2}}$, $0 \leqslant \varepsilon < 1$
Parabola	$y^2 = 4fx$	$r(\theta) = \frac{2f}{1+\cos(\theta)}$	$\varepsilon = 1$
Hyperbola	$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$	$r(\theta) = \frac{a(\varepsilon^2-1)}{1+\varepsilon\cos(\theta)}$	$\varepsilon = \sqrt{1 + \frac{b^2}{a^2}}$, $1 < \varepsilon < \infty$

Table 1.5: The four conic sections and their eccentricity parameters.

leads one of the two bodies to follow a trajectory described by one of the conic sections for which the other body is the focal point. Figure 1.76 illustrates four different trajectories for a satellite near planet F. The circle ($\varepsilon = 0$) and the ellipse ($0 \leqslant \varepsilon < 1$) describe *closed orbits*, in which the satellite is captured in the gravitational field of the planet F and remains in orbit forever. The parabola ($\varepsilon = 1$) and the hyperbola ($\varepsilon > 1$) describe *open orbits*, in which the satellite swings by the planet F and then continues.
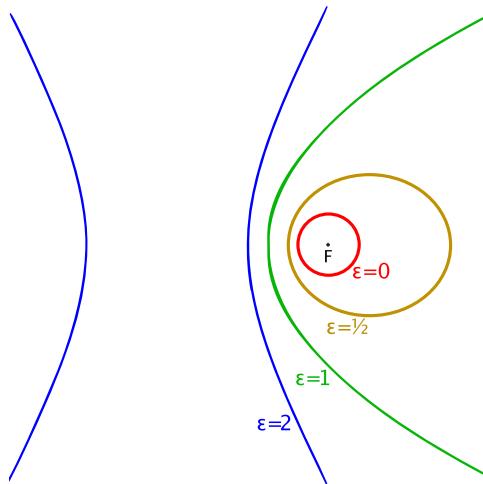
Figure 1.76: Four different trajectories for a satellite moving near a planet.

Links

[Interactive graph of a hyperbola]
`https://www.desmos.com/calculator/2mnsk5o8vn`

Discussion

The repeated use of the three algebraic techniques presented in this section will allow you to solve any system of n linear equations in n unknowns. Each time you eliminate one variable using a substitution, a subtraction, or an elimination by equating, you're simplifying the problem to a problem of finding $(n-1)$ unknowns in a system of $(n-1)$ equations. There is an entire math course called linear algebra, in which you'll develop a systematic approach for solving systems of linear equations.

Geometric solution

Solving a system of two linear equations in two unknowns can be understood geometrically as finding the point of intersection between two lines in the Cartesian plane. In this section we'll explore this correspondence between algebra and geometry to develop yet another way of solving systems of linear equations.

The algebraic equation $ax + by = c$ containing the unknowns x and y can be interpreted as a *constraint* equation on the set of possible values for the variables x and y. We can visualize this constraint geometrically by considering the coordinate pairs (x, y) that lie in the Cartesian plane. Recall that every point in the Cartesian plane can be represented as a coordinate pair (x, y), where x and y are the coordinates of the point.

Figure 1.77 shows the geometrical representation of three equations. The line ℓ_a corresponds to the set of points (x, y) that satisfy the equation $x = 1$, the line ℓ_b is the set of points (x, y) that satisfy the equation $y = 2$, and the line ℓ_c corresponds to the set of points that satisfy $x + 2y = 2$.

(a) $x = 1$ (b) $y = 2$ (c) $x + 2y = 2$
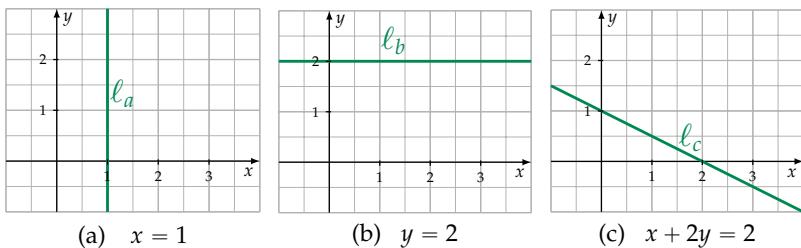
Figure 1.77: Graphical representations of three linear equations.

You can convince yourself that the geometric lines shown in Figure 1.77 are equivalent to the algebraic equations by considering individual points (x, y) in the plane. For example, the points $(1, 0)$,

Exercises

E1.39 Studious Jack borrowed $40 000 to complete his university studies and made no payments since graduation. Calculate how much money he owes after 10 years in each of the scenarios.

   a) Nominal annual interest rate of 3% compounded monthly

   b) Effective annual interest rate of 4%

   c) Nominal annual interest rate of 5% with infinite compounding

E1.40 Entrepreneurial Kate borrowed $20 000 to start a business. Initially her loan had an effective annual percentage rate of 6%, but after five years she negotiated with the bank to obtain a lower rate of 4%. How much money does she owe after 10 years?

1.23   Set notation

A *set* is the mathematically precise notion for describing a group of objects. You don't need to know about sets to perform simple math; but more advanced topics require an understanding of what sets are and how to denote set membership, set operations, and set containment relations. This section introduces all the relevant concepts.

Definitions

   - *set*: a collection of mathematical objects
   - S, T: the usual variable names for sets
   - $s \in S$: this statement is read "s is an element of S" or "s is in S"
   - $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$: some important number sets: the naturals, the integers, the rationals, and the real numbers, respectively.
   - $\varnothing$: the *empty set* is a set that contains no elements
   - { ..... }: the curly brackets are used to define sets, and the expression inside the curly brackets describes the set contents.

Set operations:

   - $S \cup T$: the *union* of two sets. The union of S and T corresponds to the elements in either S or T.
   - $S \cap T$: the *intersection* of the two sets. The intersection of S and T corresponds to the elements that are in both S and T.
   - $S \backslash T$: *set difference* or *set minus*. The set difference $S \backslash T$ corresponds to the elements of S that are not in T.

Set relations:

- ⊂: is a strict subset of
- ⊆: is a subset of or equal to

Here is a list of special mathematical shorthand symbols and their corresponding meanings:

- ∈: element of
- ∉: not an element of
- ∀: for all
- ∃: there exists
- ∄: there doesn't exist
- |: such that

These symbols are used in math proofs because they allow us to express complex mathematical arguments succinctly and precisely.

An *interval* is a subset of the real line. We denote an interval by specifying its endpoints and surrounding them with either square brackets "[" or round brackets "(" to indicate whether or not the corresponding endpoint is included in the interval.

- $[a, b]$: the *closed* interval from a to b. This corresponds to the set of numbers between a and b on the real line, including the endpoints a and b. $[a, b] = \{x \in \mathbb{R} \mid a \leqslant x \leqslant b\}$.
- (a, b): the *open* interval from a to b. This corresponds to the set of numbers between a and b on the real line, *not* including the endpoints a and b. $(a, b) = \{x \in \mathbb{R} \mid a < x < b\}$.
- $[a, b)$: the half-open interval that includes the left endpoint a but not the right endpoint b. $[a, b) = \{x \in \mathbb{R} \mid a \leqslant x < b\}$.

Sometimes we encounter intervals that consist of two disjointed parts. We use the notation $[a, b] \cup [c, d]$ to denote the union of the two intervals, which is the set of numbers *either* between a and b (inclusive) *or* between c and d (inclusive).

Sets

Much of math's power comes from *abstraction*: the ability to see the bigger picture and think *meta* thoughts about the common relationships between math objects. We can think of individual numbers like 3, -5, and π, or we can talk about the set of *all* numbers.

It is often useful to restrict our attention to a specific *subset* of the numbers as in the following examples.

always contained within the parent set, so we also know $E \cup O \subseteq \mathbb{Z}$. Combining these facts, we can establish the equality $E \cup O = \mathbb{Z}$, which states the fact, "The combination of all even and odd numbers is the same as all integers."
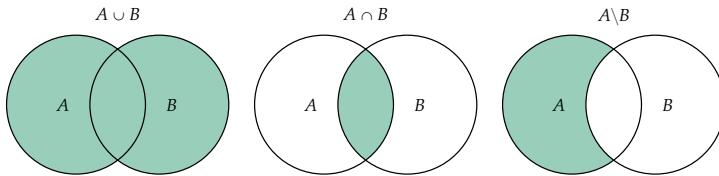
Figure 1.84: Venn diagrams showing different subsets obtained using the set operations: set union $A \cup B$, set intersection $A \cap B$, and set difference $A\backslash B$.

The set intersection $A \cap B$ and set difference $A\backslash B$ are also illustrated in Figure 1.84. The intersection of two sets contains the elements that are part of both sets. The set difference $A\backslash B$ contains all the elements that are in A but not in B.

Note the meaning of the conjunction "or" in English is ambiguous. The expression "in A or B" could be interpreted as either an "inclusive or," meaning "in A or B, or in both"—or as an "exclusive or," meaning "in A or B, but not both." Mathematicians always use "or" in the inclusive sense, so $A \cup B$ denotes elements that are in A or B, or in both sets. We can obtain an expression that corresponds to the "exclusive or" of two sets by taking the union of the sets and subtracting their intersection: $(A \cup B)\backslash(A \cap B)$.

Example 3: Set operations

Consider the three sets $A = \{a, b, c\}$, $B = \{b, c, d\}$, and $C = \{c, d, e\}$. Using set operations, we can define new sets, such as

$$A \cup B = \{a, b, c, d\}, \quad A \cap B = \{b, c\}, \quad \text{and} \quad A\backslash B = \{a\},$$

which correspond to elements in either A or B, the set of elements in A and B, and the set of elements in A but not in B, respectively.

We can also construct expressions involving three sets:

$$A \cup B \cup C = \{a, b, c, d, e\}, \qquad A \cap B \cap C = \{c\}.$$

And we can write more elaborate set expressions, like

$$(A \cup B)\backslash C = \{a, b\},$$

which is the set of elements that are in A or B but not in C.

Another example of a complicated set expression is

$$(A \cap B) \cup (B \cap C) = \{b, c, d\},$$

Sets as solutions to equations

Another context where sets come up is when describing solutions to equations and inequalities. In Section 1.1 we learned how to solve for the unknown x in equations. To solve the equation $f(x) = c$ is to find all the values of x that satisfy this equation. For simple equations like $x - 3 = 6$, the solution is a single number $x = 9$, but more complex equations can have multiple solutions. For example, the solution to the equation $x^2 = 4$ is the set $\{-2, 2\}$, since both $x = -2$ and $x = 2$ satisfy the equation.

Please update your definition of the math verb "to solve" (an equation) to include the new notion of a *solution set*—the set of values that satisfy the equation. A solution set is the mathematically precise way to describe an equation's solutions:

- The solution set to the equation $x - 3 = 6$ is the set $\{9\}$.
- The solution set for the equation $x^2 = 4$ is the set $\{-2, 2\}$.
- The solution set of $\sin(x) = 0$ is the set $\{x \mid x = \pi n, \forall n \in \mathbb{Z}\}$.
- The solution set for the equation $\sin(x) = 2$ is $\varnothing$ (the empty set), since there is no number x that satisfies the equation.

The SymPy function solve returns the solutions of equations as a list. To solve the equation $f(x) = c$ using SymPy, we first rewrite it as expression that equals zero $f(x) - c = 0$, then call the function solve:

```
>>> solve(x-3 -6, x)        # usage: solve(expr, var)
[9]

>>> solve(x**2 -4, x)
[-2, 2]

>>> solve(sin(x), x)
[0, pi]                      # found only solutions in [0,2*pi)

>>> solve(sin(x) -2, x)
[]                           # empty list = empty set
```

In the next section we'll learn how the notion of a solution set is used for describing the solutions to systems of equations.

Solution sets to systems of equations

Let's revisit what we learned in Section 1.21 about the solutions to systems of linear equations, and define their solution sets more precisely. The solution set for the system of equations

$$a_1 x + b_1 y = c_1,$$
$$a_2 x + b_2 y = c_2,$$

corresponds to the intersection of two sets:

$$\underbrace{\{(x,y) \in \mathbb{R}^2 \mid a_1x + b_1y = c_1\}}_{\ell_1} \cap \underbrace{\{(x,y) \in \mathbb{R}^2 \mid a_2x + b_2y = c_2\}}_{\ell_2}.$$

Recall that the lines ℓ_1 and ℓ_2 are the geometric interpretation of these sets. Each line corresponds to a set of coordinate pairs (x, y) that satisfy the equation of the line. The solution to the system of equations is the set of points at the intersection of the two lines $\ell_1 \cap \ell_2$. Note the word *intersection* is used in two different mathematical contexts: the solution is the *set intersection* of two sets, and also the *geometric intersection* of two lines.

Let's take advantage of this correspondence between set intersections and geometric line intersections to understand the solutions to systems of equations in a little more detail. In the next three sections, we'll look at three possible cases that can occur when trying to solve a system of two linear equations in two unknowns. So far we've only discussed Case A, which occurs when the two lines intersect at a point, as in the example shown in Figure 1.85. To fully understand the possible solutions to a system of equations, we need to think about all other cases; like Case B when $\ell_1 \cap \ell_2 = \varnothing$ as in Figure 1.86, and Case C when $\ell_1 \cap \ell_2 = \ell_1 = \ell_2$ as in Figure 1.87.

Case A: One solution. When the lines ℓ_1 and ℓ_2 are non-parallel, they will intersect at a point as shown in Figure 1.85. In this case, the solution set to the system of equations contains a single point:

$$\{(x,y) \in \mathbb{R}^2 \mid x + 2y = 2\} \cap \{(x,y) \in \mathbb{R}^2 \mid x = 1\} = \{(1, \tfrac{1}{2})\}.$$
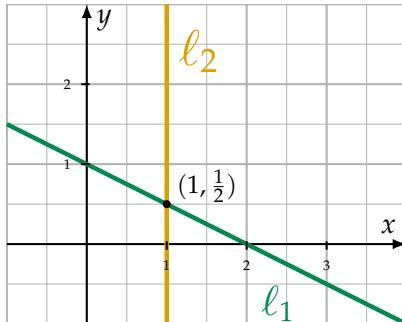
Figure 1.85: Case A: The intersection of the lines with equations $x + 2y = 2$ and $x = 1$ is the point $(1, \tfrac{1}{2}) \in \mathbb{R}^2$.

1.24   Math problems

We've now reached the first section of problems in this book. The purpose of these problems is to give you a way to comprehensively practice your math fundamentals. Knowing how to solve math problems is a very useful skill to develop. At times, honing your math chops might seem like tough mental work, but at the end of each problem, you'll gain a stronger foothold on all the topics you've been learning about. You'll also experience a small *achievement buzz* after each problem you vanquish.

Sit down and take a crack at these practice problems today, or another time when you're properly caffeinated. If you make time for some math practice, you'll develop long-lasting comprehension and true math fluency.

Without solving any problems, you're likely to forget most of what you've learned in the next few months. You might still remember the big ideas, but the details will be fuzzy and faded. By solving some of the practice problems, you'll remember a lot more stuff. Don't break the pace now: with math, it's very much *use it or lose it*!

Make sure you put your phone away while you're working on the problems. You don't need fancy technology to do math; grab a pen and some paper from the printer and you'll be fine. The great mathematicians like Descartes, Hilbert, Leibniz, and Noether did most of their work with pen and paper and they did well. Spend some time with math the way they did.

P1.1 Solve for x in the equation $x^2 - 9 = 7$.

P1.2 Solve for x in the equation $\cos^{-1}\left(\frac{x}{A}\right) - \phi = \omega t$.

P1.3 Solve for x in the equation $\frac{1}{x} = \frac{1}{a} + \frac{1}{b}$.

P1.4 Use a calculator to find the values of the following expressions:

    a) $\sqrt[4]{3^3}$ **b)** 2^{10} **c)** $7^{\frac{1}{4}} - 10$ **d)** $\frac{1}{2}\ln(e^{22})$

P1.5 Compute the following expressions involving fractions:

    a) $\dfrac{1}{2} + \dfrac{1}{4}$ **b)** $\dfrac{4}{7} - \dfrac{23}{5}$ **c)** $1\frac{3}{4} + 1\frac{31}{32}$

P1.6 Use the basic rules of algebra to simplify the following expressions:

    a) $ab\,\dfrac{1}{a}b^2cb^{-3}$ **b)** $\dfrac{abc}{bca}$ **c)** $\dfrac{27a^2}{\sqrt{9abba}}$

    d) $\dfrac{a(b+c) - ca}{b}$ **e)** $\dfrac{a}{c\sqrt[3]{b}}\dfrac{b^{\frac{4}{3}}}{a^2}$ **f)** $(x+a)(x+b) - x(a+b)$

Chapter 2

Introduction to physics

2.1 Introduction

One of the coolest things about understanding math is that you will automatically start to understand the laws of physics too. Indeed, most physics laws are expressed as mathematical equations. If you know how to manipulate equations and you know how to solve for the unknowns in them, then you know half of physics already.

Ever since Newton figured out the whole $F = ma$ thing, people have used mechanics to achieve great technological feats, like landing spaceships on the Moon. You can be part of this science thing too. Learning physics will give you the following superpowers:

1. The power to **predict the future motion of objects** using equations. For most types of motion, it is possible to find an equation that describes the position of an object as a function of time $x(t)$. You can use this equation to predict the position of the object at all times t, including the future. "Yo G! Where's the particle going to be at $t = 1.3$ seconds?" you are asked. "It is going to be at $x(1.3)$ metres." Simple as that. The equation $x(t)$ describes the object's position for *all* times t during the motion. Knowing this, you can plug $t = 1.3$ seconds into $x(t)$ to find the object's location at that time.

2. Special **physics vision** for seeing the world. After learning physics, you will start to think in terms of concepts like force, acceleration, and velocity. You can use these concepts to precisely describe all aspects of the motion of objects. Without physics vision, when you throw a ball into the air you will see it go up, reach the top, then fall down. Not very exciting. Now *with* physics vision, you will see that at $t = 0[s]$, the

same ball is thrown in the positive y-direction with an initial velocity of $v_i = 12$[m/s]. The ball reaches a maximum height of $y_{max} = \frac{12^2}{2\times 9.81} = 7.3$[m] at $t = 12/9.81 = 1.22$[s], then hits the ground after a total flight time of $t_f = 2\sqrt{\frac{2\times 7.3}{9.81}} = 2.44$[s].

The *measurement units* of physical quantities throughout this book are denoted in square brackets, like in the example above. Learning about the different measurement units is an important aspect of *physics vision*.

Why learn physics?

The main reason why you should learn physics is for the *knowledge buzz*. You will learn how to calculate the motion of objects, predict the outcomes of collisions, describe oscillations, and many other useful things. As you develop your physics skills, you will be able to use physics equations to derive one physical quantity from another. For example, we can predict the maximum height reached by a ball, if we know its initial velocity when thrown. The equations of physics are a lot like LEGO blocks; your job is to figure out different ways to connect them together.

By learning how to solve complicated physics problems, you will develop your analytical skills. Later on, you can apply these skills to other areas of life. Even if you don't go on to study science, the expertise you develop in solving physics problems will help you tackle complicated problems in general. As proof of this statement, consider the fact that companies like to hire physicists even for positions unrelated to physics: they feel confident that candidates who understand physics will be able to figure out all the business stuff easily.

Intro to science

Perhaps the most important reason you should learn physics is because it represents the golden standard for the scientific method. First of all, physics deals only with concrete things that can be **measured**. There are no feelings or subjectivities in physics. Physicists must derive mathematical models that **accurately describe** and **predict** the outcomes of experiments. Above all, we can **test** the validity of the physical models by running experiments and comparing the predicted outcome with what actually happens in the lab.

The key ingredient in scientific thinking is skepticism. Scientists must convince their peers that their equations are true without a doubt. The peers shouldn't need to *trust* the scientist; rather, they

can carry out their own tests to see if the equation accurately predicts what happens in the real world. For example, let's say I claim that the height of a ball thrown up in the air with speed 12[m/s] is described by the equation $y_c(t) = \frac{1}{2}(-9.81)t^2 + 12t + 0$. To test whether this equation is true, you can perform a throwing-the-ball-in-the-air experiment and record the motion of the ball as a video. You can then compare the motion parameters observed in the video with those predicted by the claimed equation $y_c(t)$.

- **Maximum height reached**   One thing you can check is whether the equation $y_c(t)$ predicts the ball's maximum height y_{max}. The claimed equation predicts the ball will reach its maximum height at $t = 1.22$[s]. The maximum height predicted is $\max_t\{y_c(t)\} = y_c(1.22) = 7.3$[m]. You can compare this value with the maximum height y_{max} you observe in the video.

- **Total time of flight**   You can also check whether the equation $y_c(t)$ correctly predicts the time when the ball will fall back to the ground. Using the video, suppose you measure the time it took the ball to fall back to the ground to be $t_{fall} = 2.44$[s]. If the equation $y_c(t)$ is correct, it should predict a height of zero metres for the time t_{fall}.

If both predictions of the equation $y_c(t)$ match your observations from the video, you can start to believe the claimed equation of motion $y_c(t)$ is truly an accurate model for the real world.

The scientific method depends on this interplay between experiment and theory. Theoreticians prove theorems and derive equations, while experimentalists test the validity of equations. The equations that accurately predict the laws of nature are kept while inaccurate models are rejected. At the same time, experimentalists constantly measure new data and challenge theoreticians to come up with equations that correctly describe the new measurements.

Equations of physics

The best physics equations are collected in textbooks. Physics textbooks contain only equations that have been extensively tested and are believed to be true. Good physics textbooks also explain how the equations are *derived* from first principles. This is important, because it is much easier to understand a few fundamental principles of physics, rather than memorize a long list of formulas. Understanding trumps memorization any day of the week.

The next section will teach you about three equations that fully describe the motion of any object: $x(t)$, $v(t)$, and $a(t)$. Using these equations and the equation-solving techniques from Chapter 1, we

can predict pretty much anything we want about the position and velocity of objects undergoing *constant acceleration*.

Instead of asking you to memorize these equations, I'll show you a cool trick for obtaining one equation of motion from another. These three equations describe different aspects of the same motion, so it's no surprise the equations are related. While you are not required to know how to derive the equations of physics, you do need to know how to use all these equations. Learning a bit of theory is a good deal: just a few pages of "difficult" theory (integrals) will give you a deep understanding of the relationship between $a(t)$, $v(t)$, and $x(t)$. This way, you can rely on your newly expanded math knowledge, rather than remember three separate formulas!

2.2   Kinematics

Kinematics (from the Greek word *kinema* for *motion*) is the study of trajectories of moving objects. The equations of kinematics can be used to calculate how long a ball thrown upward will stay in the air, or to calculate the acceleration needed to go from 0 to 100[km/h] in 5 seconds. To carry out these calculations, we need to choose the right *equation of motion* and figure out the values of the *initial conditions* (the initial position x_i and the initial velocity v_i). Afterward, we plug the known values into the appropriate equation of motion and solve for the unknown using one or two simple algebra steps. This entire section boils down to three equations and the plug-number-into-equation skill.
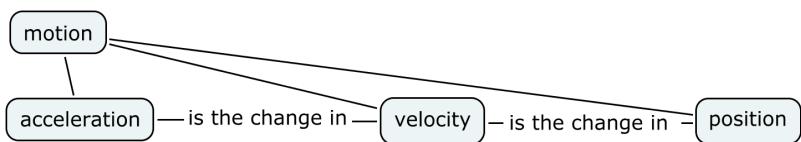

Figure 2.1: The motion of an object is described by its position, velocity, and acceleration functions.

This section is here to teach you how to use the equations of motion and help you understand the concepts of velocity and acceleration. You'll also learn how to recognize which equations to use when solving different types of physics problems.

Concepts

The key notions for describing the motion of objects are:

- t: the time. Time is measured in seconds [s].
- $x(t)$: an object's position as a function of time—also known as the equation of motion. Position is measured in metres [m] and depends on the time t.
- $v(t)$: the object's velocity as a function of time. Velocity is measured in metres per second [m/s].
- $a(t)$: the object's acceleration as a function of time. Acceleration is measured in metres per second squared [m/s^2].
- $x_i = x(0), v_i = v(0)$: the object's initial position and velocity, as measured at $t = 0$. Together x_i and v_i are known as the *initial conditions*.

Position, velocity, and acceleration

The motion of an object is characterized by three functions: the position function $x(t)$, the velocity function $v(t)$, and the acceleration function $a(t)$. The functions $x(t)$, $v(t)$, and $a(t)$ are connected—they all describe different aspects of the same motion.

You are already familiar with these notions from your experience of riding in a car. The equation of motion $x(t)$ describes the position of the car as a function of time. The *velocity* describes the change in the position of the car, or mathematically,

$$v(t) = \text{rate of change in } x(t).$$

If we measure x in metres [m] and time t in seconds [s], then the units of $v(t)$ will be metres per second [m/s]. For example, an object moving with a constant velocity of $+30$[m/s] will increase its position by 30[m] each second. Note that the velocity $v(t)$ could be positive or negative. The *speed* of an object is defined as the absolute value of its velocity $|v(t)|$. The speed is always a nonnegative quantity, whereas velocity is positive or negative depending on the direction of motion.

The rate of change of an object's velocity is called *acceleration*:

$$a(t) = \text{rate of change in } v(t).$$

Acceleration is measured in metres per second squared [m/s^2]. A constant positive acceleration means the velocity of the motion is steadily increasing, similar to pressing the gas pedal. A constant negative acceleration means the velocity is steadily decreasing, similar to pressing the brake pedal.
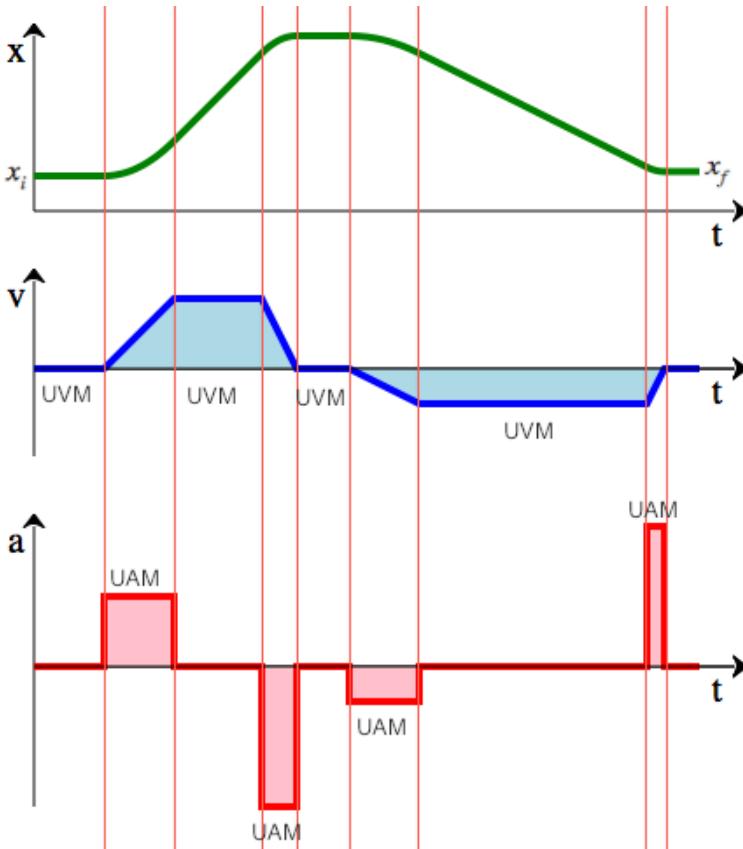
Figure 2.2: The illustration shows the simultaneous graphs of the position, velocity, and acceleration of a car during some time interval. The car starts from an initial position x_i where it sits still for some time. The driver then floors the pedal to produce a maximum acceleration for some time, and the car picks up speed. The driver then eases off the accelerator, keeping it pressed enough to maintain a constant speed. Suddenly the driver sees a police vehicle in the distance and slams on the brakes (negative acceleration) and shortly afterward brings the car to a stop. The driver waits for a few seconds to make sure the cops have passed. Next, the driver switches into reverse gear and adds gas. The car accelerates backward for a bit, then maintains a constant backward speed for an extended period of time. Note how "moving backward" corresponds to negative velocity. Finally, the driver slams on the brakes and stops the car. Notice that braking corresponds to positive acceleration when the motion is in the negative direction. The car's final position is x_f.

In a couple of paragraphs, we'll discuss the exact mathematical equations for $x(t)$, $v(t)$, and $a(t)$, but before we dig into the math, let's look at the example of the motion of a car illustrated in Figure 2.2. We can observe two distinct types of motion. During some times, the car undergoes motion at a constant velocity (uniform velocity motion, UVM). During other times, the car undergoes motion with constant acceleration (uniform acceleration motion, UAM). There exist many other types of motion, but for the purpose of this section we'll focus on these two types of motion.

- UVM: During times when there is no acceleration, the car maintains a uniform velocity and therefore $v(t)$ is a constant function. For motion with constant velocity, the position function is a line with a constant slope because, by definition, $v(t) = $ slope of $x(t)$.

- UAM: During times where the car experiences a constant acceleration $a(t) = a$, the velocity of the function changes at a constant rate. The rate of change of the velocity is constant $a = $ slope of $v(t)$, so the velocity function looks like a line with slope a. The position function $x(t)$ has a curved shape (quadratic) during moments of constant acceleration.

Formulas

There are basically four equations you need to know for this entire section. Together, these four equations fully describe all aspects of motion with constant acceleration.

Uniformly accelerated motion (UAM)

If the object undergoes a *constant* acceleration $a(t) = a$—like a car when you floor the *accelerator*—then its motion can be described by the following equations:

$$a(t) = a, \tag{2.1}$$
$$v(t) = at + v_i, \tag{2.2}$$
$$x(t) = \tfrac{1}{2}at^2 + v_i t + x_i, \tag{2.3}$$

where v_i is the initial velocity of the object and x_i is its initial position.

Here is another useful equation to remember:

$$[v(t)]^2 = v_i^2 + 2a[x(t) - x_i],$$

which is usually written

$$v_f^2 = v_i^2 + 2a\Delta x, \tag{2.4}$$

where v_f denotes the final velocity (at $t = t_f$) and Δx denotes the *change* in the x-coordinate between $t = 0$ and $t = t_f$. The triangle thing Δ is the capital Greek letter *delta*, which is often used to denote the change in quantities. Using the Δ-notation, we can rewrite equation (2.2) as follows: $\Delta v = a\Delta t$, where $\Delta v = v_f - v_i$ and $\Delta t = t_f - t_i$.

That's it! Memorize these four equations, plug-in the right numbers, and you can solve any kinematics problem humanly imaginable.

Uniform velocity motion (UVM)

The special case where there is zero acceleration ($a = 0$), is called *uniform velocity motion* or UVM. The velocity stays uniform (constant) because there is no acceleration. The following three equations describe the motion of an object with uniform velocity:

$$a(t) = 0,$$
$$v(t) = v_i,$$
$$x(t) = v_i t + x_i.$$

As you can see, these are really the same equations as in the UAM case above, but because $a = 0$, some terms are missing.

Free fall

We say an object is in *free fall* if the only force acting on it is the force of gravity. On the surface of the Earth, the force of gravity produces a constant acceleration of $a_y = -9.81[\text{m/s}^2]$. The negative sign is there because the gravitational acceleration is directed downward, and we assume the y-axis points upward. Since the gravitational acceleration is constant, we can use the UAM equations to find the height $y(t)$ and velocity $v(t)$ of objects in free fall.

Examples

Now we'll illustrate how the equations of kinematics are used.

Moroccan example   Suppose your friend wants to send you a ball wrapped in aluminum foil by dropping it from his balcony, which is located at a height of $y_i = 44.145[\text{m}]$. How long will it take for the ball to hit the ground?

We recognize this is a problem with acceleration, so we start by writing the general UAM equations:

$$y(t) = \tfrac{1}{2}at^2 + v_i t + y_i,$$
$$v(t) = at + v_i.$$

To find the answer, substitute the following known values into the $y(t)$ equation: $y(0) = y_i = 44.145$[m]; $a = -9.81$ (since the ball is in free fall); and $v_i = 0$[m/s] (since the ball was released from rest). We want to find the time t_{fall} when the height of the ball will be zero:

$$0 = y(t_{\text{fall}}),$$
$$0 = \tfrac{1}{2}(-9.81)(t_{\text{fall}})^2 + 0(t_{\text{fall}}) + 44.145.$$

Solving for t_{fall} we find the answer $t_{\text{fall}} = \sqrt{\frac{44.145 \times 2}{9.81}} = 3$[s].

As another variation of this type of kinematics question, suppose you're given the time it takes for the ball to fall $t_{\text{fall}} = 3$[s], and you're asked to find the height of the balcony. You already know $y(3) = 0$, and are looking for the initial height y_i. You can solve for y_i in the equation $0 = \tfrac{1}{2}(-9.81)3^2 + y_i$. The answer gives $y_i = 44.145$[m].

0 to 100 in 5 seconds   Say you're in the driver's seat of a car and you want to accelerate from 0 to 100[km/h] in 5 seconds. How much acceleration must the car's engine produce, assuming it produces a constant amount of acceleration?

We can calculate the necessary a by plugging the required values into the velocity equation for UAM:

$$v(t) = at + v_i.$$

Before we tackle that, we need to convert the velocity in [km/h] to velocity in [m/s]: 100[km/h] $= \frac{100[\text{km}]}{1[\text{h}]} \cdot \frac{1000[\text{m}]}{1[\text{km}]} \cdot \frac{1[\text{h}]}{3600[\text{s}]} = 27.8$ [m/s]. We substitute the desired values $v_f = 27.8$[m/s], $v_i = 0$, and $t = 5$[s] into the equation for $v(t)$ and solve for a:

$$27.8 = v(5) = a5 + 0.$$

After solving for a, we find the car's engine must produce a constant acceleration of $a = \frac{27.8}{5} = 5.56$[m/s^2] or greater.

Moroccan example II   Some time later, your friend wants to send you another aluminum ball from his apartment located on the 14th floor (height of 44.145[m]). To decrease the time of flight, he *throws*

the ball straight down with an initial velocity of 10[m/s]. How long does it take for the ball to hit the ground?

Imagine the apartment building as a y-axis that measures distance upward starting from the ground floor. We know the balcony is located at a height of $y_i = 44.145$[m], and that at $t = 0$[s] the ball starts with $v_i = -10$[m/s]. The initial velocity is negative because it points in the opposite direction of the y-axis. We also know there is an acceleration due to gravity of $a_y = -9.81$[m/s^2].

We start by writing the general UAM equation:

$$y(t) = \tfrac{1}{2}a_y t^2 + v_i t + y_i.$$

To find the time when the ball will hit the ground, we can solve for t in the equation $y(t) = 0$, then plug the values into the UAM equation,

$$y(t) = 0 = \tfrac{1}{2}(-9.81)t^2 - 10t + 44.145,$$

and solve for t using the quadratic formula. First, rewrite the quadratic equation in standard form: $0 = \underbrace{4.905}_{a}\, t^2 \underbrace{+10.0}_{b}\, t \underbrace{-44.145}_{c}$.

Then solve using the quadratic equation:

$$t_{\text{fall}} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a} = \frac{-10 \pm \sqrt{100 + 866.12}}{9.81} = 2.15 \qquad [\text{s}].$$

We ignore the negative-time solution because it corresponds to a time in the past. Compared to the first Moroccan example, we see that throwing the ball downward makes it fall to the ground in less time.

Discussion

Most kinematics problems you'll be solving follow the same pattern as the examples above. Given some initial values, you'll be asked to solve for some unknown quantity.

It's important to keep in mind the *signs* of the numbers you plug into the equations. You should always draw the coordinate system and indicate clearly (to yourself) the x-axis, which measures the object's displacement. A velocity or acceleration quantity that points in the same direction as the x-axis is a positive number, while quantities pointing in the opposite direction are negative numbers.

By the way, all this talk about $v(t)$ being the "rate of change of $x(t)$" is starting to get on my nerves. The expression "rate of change of" is an indirect way of saying the calculus term *derivative*. In order to use this more precise terminology throughout the remainder of the book, we'll now take a short excursion into the land of calculus to define two fundamental concepts: derivatives and integrals.

Exercises

E2.1 Calculate the time it will take a rocket launched with initial velocity v_i[m/s] and constant acceleration of a[m/s^2] to reach the velocity of 100[m/s]. What is the total distance travelled?

a) $v_i = 20$[m/s], $a = 5$[m/s^2]   **b)** $v_i = 30$[m/s], $a = 10$[m/s^2]

2.3 Introduction to calculus

Calculus is the study of functions and their properties. The two operations in the study of calculus are derivatives—which describe how quantities *change over time*—and integrals, which are used to calculate the total amount of a quantity *accumulated* over a time period.

Derivatives

The derivative function $f'(t)$ describes how the function $f(t)$ changes over time. The derivative encodes the information about the *instantaneous rate of change* of the function $f(t)$, which is the same as the *slope* of the graph of the function at that point:

$$f'(t) = \text{slope}_f(t) = \frac{\text{change in } f(t)}{\text{change in } t} = \frac{f(t + \Delta t) - f(t)}{\Delta t}.$$

If the derivative $f'(t)$ is equal to 5 units per second, this means that $f(t)$ changes by 5 units each second. The derivative of the constant function is zero because it has zero rise-over-run everywhere. The derivative of the function $f(t) = mt + b$ (a line) is the constant function $f'(t) = m$. More generally, the instantaneous slope of a function is different for different values of t, as illustrated in Figure 2.3.
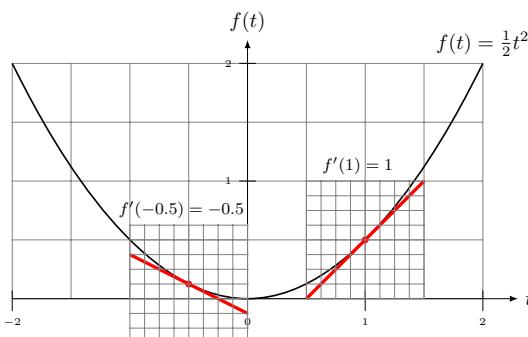

Figure 2.3: The derivative of the function at $t = a$ is denoted $f'(a)$ and describes the slope function at that point.

The derivative operation is denoted by several symbols: $Df(t) = f'(t) = \frac{df}{dt} = \frac{d}{dt}\{f(t)\} = \dot{f}$. All these symbols carry the same meaning. Think of $f'(t)$ not as a separate entity from $f(t)$, but as a *property* of the function $f(t)$. It's best to think of the derivative as an *operator* $\frac{d}{dt}$ that you apply to functions to obtain their slope information.

Integrals

An integral corresponds to the computation of the *area* enclosed between the curve $f(t)$ and the t-axis over some interval:

$$A(a,b) = \int_{t=a}^{t=b} f(t)\, dt.$$

The symbol $\int$ is shorthand for *sum*. Indeed, the area under the curve corresponds to the sum of the values of the function $f(t)$ between $t = a$ and $t = b$. The integral is the total of f between a and b.
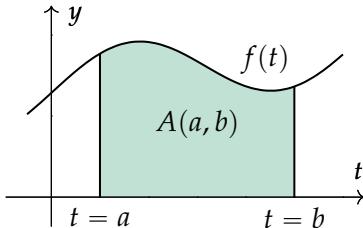
Figure 2.4: The integral of the function $f(t)$ between $t = a$ and $t = b$.

Example 1

We can easily find the area under the graph of the constant function $f(t) = 3$ between any two points because the region under the curve is rectangular. Choosing $t = 0$ as the starting point, we obtain the integral function $F(\tau)$, which corresponds to the area under $f(t)$ between $t = 0$ and $t = \tau$:

$$F(\tau) = A(0, \tau) = \int_0^\tau f(t)\, dt = 3\tau.$$

The area is equal to the rectangle's height times its width as illustrated in Figure 2.5.
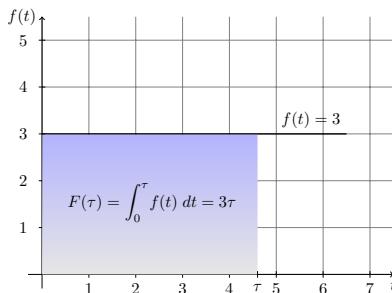
Figure 2.5: The area of a rectangle of height 3 and width τ is equal to 3τ.

Example 2

Consider now the area under the graph of the line $g(t) = t$, starting from $t = 0$. Since the region under the curve is triangular, we can compute its area. Recall the area of a triangle is given by the length of its base times its height divided by 2.

The general formula for the area under $g(t)$ from $t = 0$ until $t = \tau$ is described by the following integral calculation:

$$G(\tau) = A(0, \tau) = \int_0^\tau g(t)\, dt = \frac{\tau \times \tau}{2} = \tfrac{1}{2}\tau^2.$$
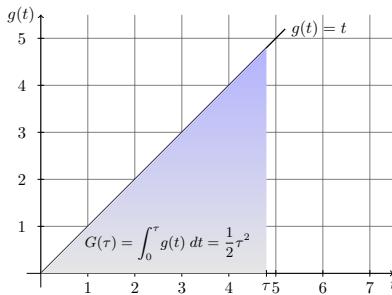
Figure 2.6: The area of a triangle with base τ and height τ is equal to $\tfrac{1}{2}\tau^2$.

We're able to compute the above integrals thanks to the simple geometries of the areas under the graphs. Later in this book (Chapter 5), we'll develop techniques for finding integrals (areas under the curve) of more complicated functions. In fact, there is an entire course called integral calculus, which is dedicated to the task of finding integrals.

But don't worry, you don't need to know *everything* about integrals to learn physics. What is important right now is that you understand the concept of integration. The integral of a function is the area under the graph of the function, which is in some sense the total amount of the function accumulated during some interval of time. For the most part of the first-year physics, the only integral formulas you'll need to know are

$$\int_0^\tau a\ dt = a\tau \qquad \text{and} \qquad \int_0^\tau at\ dt = \tfrac{1}{2}a\tau^2.$$

The first integral describes the general calculation of the area under a constant function, like in Example 1. The second formula is a generalization of the formula we derived in Example 2. Using these formulas in combination, you can compute the integral of an arbitrary line $h(t) = mt + b$ as follows:

$$H(\tau) = \int_0^\tau h(t)\ dt = \int_0^\tau (mt + b)\ dt = \int_0^\tau mt\ dt\ + \int_0^\tau b\ dt = \tfrac{1}{2}m\tau^2 + b\tau.$$

Regroup

At this point you might be on the fence about the new calculus concepts. On one hand, calculating slopes (derivatives) and areas under the curve (integrals) seem like trivial tasks. On the other hand, seeing five different notations for the derivative and the weird integral sign has probably put some fear in you. You might be wondering whether you *really* need to learn about derivatives and integrals. How often do you have to compute the area under the graph of a function in the real world? It turns out that "calculating the area under a curve" is very useful since it is the "undo operation" for the derivative.

Inverse operations

The integral is the inverse operation of the derivative. Many equations in math and physics involve the *derivative* of some unknown function. Understanding the inverse relationship between integrals and derivatives will allow you to solve for the unknown function in these equations.

You should already be familiar with the inverse relationship between functions. When solving equations, we use inverse functions to *undo* functions that stand in our way as we try to isolate the unknown x. Similarly, we use the integral operation to *undo* the effects of the derivative operation when we try to solve for some unknown

function $f(t)$. For example, suppose $g(t)$ is a known function and we're trying to solve for $f(t)$ in the equation

$$\frac{d}{dt}\{f(t)\} = g(t).$$

Taking the integral on the left-hand side of the equation will undo the derivative operation. To keep the equality true, we must apply the "integrate over t" operation on both sides of the equation. We obtain

$$\int \frac{d}{dt}\{f(t)\}\,dt = \int g(t)\,dt,$$

$$f(t) = \int g(t)\,dt.$$

The take-home message is that every time you want to *undo* a derivative, you can apply the integral operation, however, there is a little technical detail we must clarify to make this statement precise.

The integral isn't *exactly* the inverse of the derivative—there is a tricky extra constant factor that appears when we integrate. Let's analyze in more detail what happens when we perform the combo of the derivative operation followed by the integral operation on some function $f(t)$. Suppose we are given the derivative function $f'(t)$ and asked to integrate it between $t = 0$ and $t = \tau$. Intuitively, this integral corresponds to calculating the **total of the changes** in $f(t)$ during that time interval. Recall the notation for "change in f" $\Delta f = f(\tau) - f(0)$, which we used previously. This notation makes it easy to see how the integral over $f'(t)$ corresponds to the total change in $f(t)$ between $t = 0$ and $t = \tau$:

$$\int_0^\tau f'(t)\,dt = \Delta f = f(\tau) - f(0).$$

Calculating the total of the instantaneous changes in f during the interval $[0, \tau]$ is the same as finding the change in $f(t)$ between the endpoints of the interval. If we rewrite the above equation to isolate $f(\tau)$, we obtain

$$f(\tau) = f(0) + \int_0^\tau f'(t)\,dt.$$

Note that the expression for $f(\tau)$ depends on the value of $f(t)$ at $t = 0$, which we call the *initial value* of the function. In physics problems, the initial values of the equations of motion $x_i \stackrel{\text{def}}{=} x(0)$ and $v_i \stackrel{\text{def}}{=} v(0)$ are called the *initial conditions*.

Banking example   To illustrate how derivative and integral operations apply to the real world, I'll draw an analogy from a scenario that every student is familiar with. Consider the function $ba(t)$, which represents your bank account balance at time t. Also consider the function $tr(t)$, which corresponds to the transactions (deposits and withdrawals) on your account.

The function $tr(t)$ is the derivative of the function $ba(t)$. If you ask, "How does my balance change over time?" the answer is the function $tr(t)$. Using mathematical symbols, we can represent this relationship as

$$tr(t) = \frac{d}{dt}\{ba(t)\}.$$

If the derivative is positive, your account balance is growing. If the derivative is negative, your account balance is depleting.

Suppose you have a record of all the transactions on your account $tr(t)$, and you want to compute the final account balance at the end of the month. Since $tr(t)$ is the derivative of $ba(t)$, you can use an integral (the inverse operation of the derivative) to obtain $ba(t)$. Knowing the balance of your account at the beginning of the month, you can predict the balance at the end of the month by using the following integral calculation:

$$ba(30) = ba(0) + \int_0^{30} tr(t)\, dt.$$

This calculation makes sense since $tr(t)$ represents the instantaneous changes in $ba(t)$. If you want to find the overall change in the account balance from day 0 until day 30, you can compute the total of all the transactions on the account.

We use integrals every time we need to calculate the total of some quantity over a time period. In the next section, we'll see how these integration techniques can be applied to the subject of kinematics, and how the equations of motion for UAM are derived from first principles.

2.4   Kinematics with calculus

To carry out kinematics calculations, all we need to do is plug the initial conditions (x_i and v_i) into the correct equation of motion. But how did Newton come up with the equations of motion in the first place? Now that you know Newton's mathematical techniques (calculus), you can learn to derive the equations of motion by yourself.

Concepts

Recall the kinematics concepts related to the motion of objects:

- t: time
- $x(t)$: position as a function of time
- $v(t)$: velocity as a function of time
- $a(t)$: acceleration as a function of time
- $x_i \overset{\text{def}}{=} x(0), v_i \overset{\text{def}}{=} v(0)$: the initial conditions

Position, velocity, and acceleration revisited

The equations of kinematics are used to predict the motion of objects. Suppose you know the acceleration of the object $a(t)$ at all times t. Can you find $x(t)$ starting from $a(t)$?

The equations of motion $x(t)$, $v(t)$, and $a(t)$ are related:

$$a(t) \xleftarrow{\frac{d}{dt}} v(t) \xleftarrow{\frac{d}{dt}} x(t).$$

The velocity function is the derivative of the position function and the acceleration function is the derivative of the velocity function.

General procedure

If you know the acceleration of an object as a function of time $a(t)$, and you know its initial velocity $v_i = v(0)$, you can find its velocity function $v(t)$ for all later times using integration. This is because the acceleration function $a(t)$ describes the change in the object's velocity.

Let's see how this works. The object started with an initial velocity of $v_i = v(0)$ at $t = 0$. The velocity at a later time $t = \tau$ is equal to v_i plus the total acceleration of the object between $t = 0$ and $t = \tau$:

$$v(\tau) = v_i + \int_0^\tau a(t)\, dt.$$

If you know the initial position x_i and the velocity function $v(t)$, you can find the position function $x(t)$ by using integration. We find the position at time $t = \tau$ by adding all the velocities (the changes in the object's position) that occurred between $t = 0$ and $t = \tau$:

$$x(\tau) = x_i + \int_0^\tau v(t)\, dt.$$

The procedure for finding $x(t)$ starting from $a(t)$ can be summarized as follows:

$$a(t) \xrightarrow{v_i + \int dt} v(t) \xrightarrow{x_i + \int dt} x(t).$$

Next, I'll illustrate how you can apply this procedure to the important special case of an object undergoing uniformly accelerated motion.

Derivation of the UAM equations of motion

Consider an object undergoing uniformly accelerated motion (UAM) with acceleration function $a(t) = a$. Suppose we know the initial velocity of v_i, and we want to find the velocity at a later time $t = \tau$. We compute $v(\tau)$ using the following integral:

$$v(\tau) = v_i + \int_0^\tau a(t)\, dt = v_i + \int_0^\tau a\, dt = v_i + a\tau.$$

Velocity as a function of time is given by the initial velocity v_i added to the integral of the acceleration. The integration can be visualized as the calculation of the area of a rectangle, similar to the calculation we saw in Example 1 on page 166.

You can also use integration to find the position function $x(t)$ if you know the initial position x_i and the velocity function $v(t)$. The formula is

$$x(\tau) = x_i + \int_0^\tau v(t)\, dt = x_i + \int_0^\tau (v_i + at)\, dt = x_i + v_i\tau + \tfrac{1}{2}a\tau^2.$$

The integration step can be visualized as the calculation of the area of a triangle with slope a stacked on top of a rectangle of height v_i.

Note that the above calculations required knowing the initial conditions x_i and v_i. These initial values were required because the integral calculations we performed only told us the *change* in the quantities relative to their initial values.

The fourth equation

We can derive the fourth equation of motion,

$$v_f^2 = v_i^2 + 2a(x_f - x_i),$$

by combining the equations of motion $v(t)$ and $x(t)$. Let's see how. Start by squaring both sides of the velocity equation $v_f = v_i + at$ to obtain

$$v_f^2 = (v_i + at)^2 = v_i^2 + 2av_i t + a^2 t^2 = v_i^2 + 2a[v_i t + \tfrac{1}{2}at^2].$$

The term in the square bracket is equal to $\Delta x = x(t) - x_i = x_f - x_i$.

Applications of derivatives

Recall that the velocity and the acceleration functions are obtained by taking derivatives of the position function:

$$x(t) \xrightarrow{\frac{d}{dt}} v(t) \xrightarrow{\frac{d}{dt}} a(t).$$

We just saw how to use integration to follow this chain of operations in reverse to obtain $x(t)$ for the special case of constant acceleration:

$$a(t) = a,$$

$$v(t) = v_i + \int_0^t a(\tau)\, d\tau = v_i + at,$$

$$x(t) = x_i + \int_0^t v(\tau)\, d\tau = x_i + v_i t + \tfrac{1}{2}at^2.$$

Note that, in addition to the integral calculations, the formulas for $v(t)$ and $x(t)$ require some additional information—the initial value of the function.

Earlier we defined the derivative operator $\frac{d}{dt}$ that computes the derivative function $f'(t)$, which tells us the slope of the function $f(t)$. There are several derivative formulas that you need to learn to be proficient at calculus. We'll get to them in Chapter 5. For now, the only derivative formula you'll need is the *power rule* for derivatives:

$$\text{if} \quad f(t) = At^n \quad \text{then} \quad f'(t) = nAt^{n-1}.$$

Using this formula on each term in the function $f(t) = A + Bt + Ct^2$ we find its derivative is $\frac{df}{dt} = f'(t) = 0 + B + 2Ct$.

Let's now use the derivative to verify that the equations of motion we obtained above satisfy $v'(t) = a(t)$ and $x'(t) = v(t)$. Applying the derivative operation to both sides of the equations we obtain

$$a'(t) = 0,$$

$$v'(t) = \tfrac{d}{dt}\{v_i + at\} = \tfrac{d}{dt}\cancel{\{v_i\}} + \tfrac{d}{dt}\{at\} = 0 + a \qquad = \; a(t),$$

$$x'(t) = \tfrac{d}{dt}\cancel{\{x_i\}} + \tfrac{d}{dt}\{v_i t\} + \tfrac{d}{dt}\{\tfrac{1}{2}at^2\} = 0 + v_i + at \qquad = \; v(t).$$

Note that computing the derivative of a function kills the information about its initial value; the derivative contains only information about the changes in $f(t)$.

Let's summarize what we've learned so far about derivatives and integrals. Integrals are useful because they allow us to compute $v(t)$ from $a(t)$, and $x(t)$ from $v(t)$. The derivative operation is useful

because it allows us to obtain $v(t)$ if we know $x(t)$, and/or obtain $a(t)$ if we know $v(t)$. Recall that $x(t)$, $v(t)$, and $a(t)$ correspond to three different aspects of the same motion, as shown in Figure 2.2 on page 160. The operations of calculus allow us to move freely between the different descriptions of the motion.

Discussion

According to Newton's second law of motion, forces are the cause of acceleration and the formula that governs this relationship is

$$F_{net} = ma,$$

where F_{net} is the magnitude of the net force acting on the object.

In Chapter 4 we'll learn about *dynamics*, the study of the different kinds of forces that can act on objects: gravitational force $\vec{F}_g$, spring force $\vec{F}_s$, friction force $\vec{F}_f$, and other forces. To find an object's acceleration, we must add together all the forces acting on the object and divide by the object's mass:

$$\sum F_i = F_{net} \qquad \Rightarrow \qquad a = \frac{1}{m} F_{net}.$$

The physics procedure for predicting the motion of an object given the forces acting on it can be summarized as follows:

$$\frac{1}{m} \underbrace{\left(\sum \vec{F} = \vec{F}_{net} \right)}_{\text{dynamics}} = \underbrace{a(t) \xrightarrow{v_i + \int dt} v(t) \xrightarrow{x_i + \int dt} x(t)}_{\text{kinematics}}.$$

Free fall revisited

The force of gravity acting on an object of mass m on the surface of the Earth is given by $\vec{F}_g = -mg\hat{y}$, where $g = 9.81[\text{m/s}^2]$ is the *gravitational acceleration* on the surface of the Earth. We previously discussed that an object is in *free fall* when the only force acting on it is the force of gravity. In this case, Newton's second law tells us

$$\vec{F}_{net} = m\vec{a}$$
$$-mg\hat{y} = m\vec{a}.$$

Dividing both sides by the mass, we see the acceleration of an object in free fall is $\vec{a} = -9.81\hat{y}$.

It's interesting to note that an object's mass does not affect its acceleration during free fall. The force of gravity is proportional to the

mass of the object, but acceleration is inversely proportional to the mass of the object; overall, it holds that $a_y = -g$ for objects in free fall, regardless of their mass. This observation was first made by Galileo in his famous Leaning Tower of Pisa experiment. Galileo dropped a wooden ball and a metal ball (same shape, different mass) from the Leaning Tower of Pisa, and observed that they fell to the ground at the same time. Search for "Apollo 15 feather and hammer drop" on YouTube to see this experiment performed on the Moon.

What next?

You might have noticed that in the last couple of paragraphs we started putting little arrows on top of certain quantities. The arrows are there to remind you that forces, velocities, and accelerations are *vector quantities*. In the next chapter, we'll make a short mathematical digression to introduce all the vectors concepts necessary to understand physics.
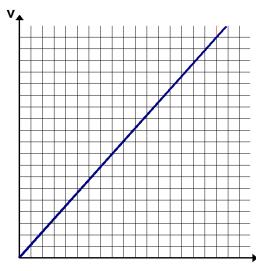
Before we proceed with the vectors lessons and more advanced physics topics, it's a good idea to practice using the physics equations you just learned. Take the time to solve a couple of the practice problems in the next section.

2.5 Kinematics problems

We spent an entire chapter learning about position, velocity, and acceleration equations used to describe the motion of objects. It's now time to practice using these equations to solve problems.

Here are some general tips for solving kinematics problems. First, try to determine which equation you'll need to solve the problem. There are just four of them: $x(t) = x_i + v_i t + \frac{1}{2}at^2$, $v(t) = v_i + at$, $a(t) = a = \frac{F_{net}}{m}$, and $v_f^2 = v_i^2 + 2a\Delta x$, so it can't be that hard. If you can't figure it out, check the hint then try solving the problem. Always draw a diagram labelling all the variables that appear in your equations. This way you'll always have a picture of what is going on. Check your answer against the answer provided on page 460. If you didn't get the right answer, check your work and try again. Don't look at the solution yet. Try to figure out the problem by revisiting the assumptions you made, the equations you wrote, and the steps you followed. Look at the solution only if you can't figure out the problem after 10 minutes and you're running out of ideas.

P2.1   Below is a velocity-vs-time graph of a moving particle. Is the parti-
cle gaining or losing speed? Does the graph describe uniformly accelerated
motion (UAM) or not?

Hint: Acceleration is the slope of the velocity graph.

P2.2   You're running away from point A. At $t = 2$[s] you're 3[m] away from
A, at $t = 4$[s] you're 8[m] away from A, and at $t = 6$[s] you're 14[m] away
from A. Are you running with uniform velocity (UVM)?
Hint: Calculate the velocity during each time interval.

P2.3   A car is moving on a straight road. Indicate whether the car's speed
is increasing or decreasing in the following cases:
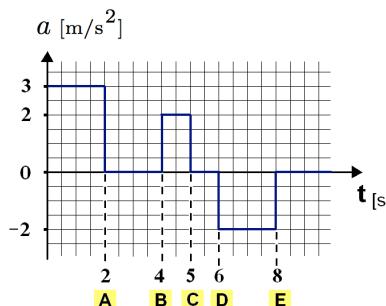
1. Velocity is negative, acceleration is positive.

2. Velocity is negative, acceleration is negative.

Hint: Pay attention to the relative direction of acceleration to velocity.

P2.4   A body is pushed by a constant force F, and at time $t = t_0$ the force be-
comes zero. Determine when the particle is in uniformly accelerated motion
(UAM) and when it is in uniform velocity motion (UVM).
Hint: Remember Newton's 2nd law of motion.

P2.5   A car has the following acceleration-vs-time graph. The car starts from
rest at $t = 0$[s]. Find the velocity of the car at times A, B, C, D, and E.

Hint: The change in velocity is the area under the acceleration graph.

P2.6   The position of a rocket as a function of time is described by the equa-
tion $x(t) = 3t^3 + 5t^2 - 3t + 5$. Find the velocity and the acceleration of the
rocket as functions of time.
Hint: Differentiate the function with respect to t.

P2.7   You're on a mission to Jupiter where you design an experiment to
measure the planet's gravitational acceleration. In the experiment, you let
go of a ball from a height of 4[m] and watch it fall to the ground. When the
ball hits the ground, its speed is 14[m/s].

1. What is the gravitational acceleration on Jupiter?

2. Find the position of the ball as a function of time.

Hint: Use the fourth equation of motion.

P2.8   You're pulling a 5[kg] cart in a straight path. The position of the cart
as a function of time is $x(t) = 6t^2 + 2t + 1$[m].

1. Find the velocity and acceleration of the cart as functions of time.

2. Calculate the force you're using to pull the cart.

Hint: Take the derivative of the position with respect to time. Use Newton's
2nd law $F = ma$.
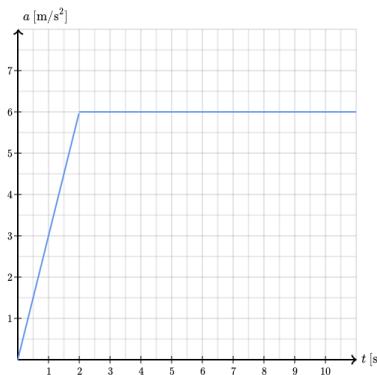
P2.9   A remote controlled car has a mass of 0.5[kg]. The electric engine
pushes the car with a force of 1.0[N] starting from rest at point A.

1. Find the acceleration, velocity, and position of the car as functions of
   time, assuming $x = 0$ at point A.

2. Calculate the velocity of the car at $t = 4$[s].

3. What is the car's velocity when it is 9[m] away from point A?

Hint: Use Newton's 2nd law and integration.

P2.10   Below is an acceleration-vs-time graph of a particle. At $t = 0$[s],
the particle starts moving from rest at $x = 0$[m]. The particle's acceleration
from $t = 0$[s] to $t = 2$[s] is given by $a(t) = 3t$[m/s²]. After $t = 2$[s], the
acceleration is constant $a = 6$[m/s²].
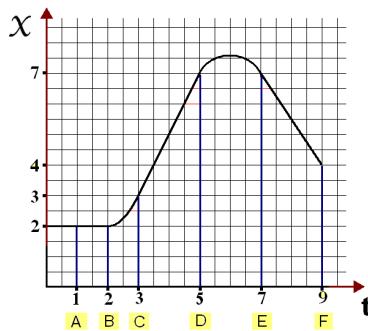
1. Find the velocity $v(2)$ and position $x(2)$ of the particle at $t = 2$[s].

2. Construct the functions of time that describe the acceleration, the ve-
   locity, and the position of the particle after $t = 2$[s].

3. How much time is needed for the particle to reach $x = 49$[m]?

4. At what distance from the origin will the particle's velocity reach
   12[m/s]?

Hint: Use integration to find the velocity and the position. The integral of $f(t) = t^2$ is $F(t) = \frac{1}{3}t^3$. Make sure that when $t = 2[s]$, the functions $v(t)$ and $x(t)$ in Part 2 match your answer from Part 1.

P2.11 The graph below shows the position-vs-time graph of a squirrel running in a field where x is in metres and t is in seconds.

1. Calculate the squirrel's velocity during the time intervals A to B, C to D, and E to F.

2. Indicate whether the squirrel is standing still, moving forward (in the positive x-direction), or moving backward during the following intervals: 0[s] to 2[s], 2[s] to 6[s], and 6[s] to 9[s].

Hint: Velocity is Δx divided by Δt.

P2.12 A car passes point A with velocity $v_i[m/s]$ at $t = 0[s]$, with an acceleration of $-2[m/s^2]$. The car comes to rest 9[m] from point A.

1. What is v_i?

2. What is the position of the car as a function of time?

Hint: Use the fourth equation of motion.

P2.13 Two dogs are running after a tennis ball. At $t = 0[s]$ the first dog starts running from rest with an acceleration of $3[m/s^2]$. The other dog is $4[m]$ ahead of the first dog at $t = 0[s]$, running with velocity $3[m/s]$ and acceleration $1[m/s^2]$.

1. Construct the positions of the two dogs as functions of time.

2. At what time will the dogs meet?

Hint: Equate the positions of both dogs.

P2.14 A car has the position function $x(t) = 2t^2 + 5t + 7[m]$.

1. What is the car's position and velocity at $t = 0[s]$?

2. Find the velocity and acceleration of the car as functions of time.

3. Find the position and the velocity of the car at $t = 5[s]$?

Hint: Use differentiation to find the velocity and acceleration functions.

P2.15 A car moving with initial velocity v_i applies the brakes. After $2[m]$ the car's speed is $4[m/s]$, and $4[m]$ after applying the brakes the car comes to a stop. What is v_i and how much time was needed for the car to stop? Assume the car's acceleration (deceleration) is constant starting from the point the brakes are applied until it stops.

Hint: Use the fourth equation of motion. Find the position function.

Chapter 3

Vectors

In this chapter we'll learn how to manipulate multi-dimensional objects called vectors. Vectors are the precise way to describe directions in space. We need vectors in order to describe physical quantities like forces, velocities, and accelerations.

Vectors are built from ordinary numbers, which form the *components* of the vector. You can think of a vector as a list of numbers, and *vector algebra* as operations performed on the numbers in the list. Vectors can also be manipulated as geometric objects, represented by arrows in space. For instance, the arrow that corresponds to the vector $\vec{v} = (v_x, v_y)$ starts at the origin $(0,0)$ and ends at the point (v_x, v_y). The word vector comes from the Latin *vehere*, which means *to carry*. Indeed, the vector $\vec{v}$ takes the point $(0,0)$ and carries it to the point (v_x, v_y).
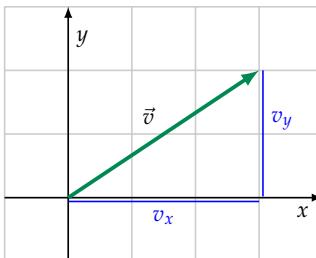
Figure 3.1: The vector $\vec{v} = (3,2)$ is an arrow in the Cartesian plane. The horizontal component of $\vec{v}$ is $v_x = 3$ and the vertical component is $v_y = 2$.

This chapter will introduce you to vectors, vector algebra, and vector operations, which are very useful for solving physics problems. What you'll learn here applies more broadly to current problems in computer graphics, probability theory, machine learning, and other

fields of science and mathematics. It's all about vectors these days, so you'd best get to know them.
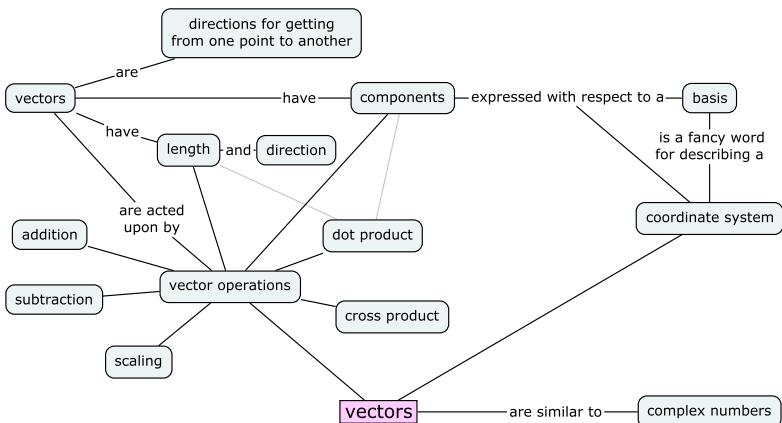

Figure 3.2: This figure illustrates the new concepts related to vectors. As you can see, there is quite a bit of new vocabulary to learn, but don't be fazed—all these terms are just fancy ways of talking about arrows.

3.1   Great outdoors

Vectors are directions for getting from point A to point B. Directions can be given in terms of street names and visual landmarks, or with respect to a coordinate system.

While on vacation in British Columbia, you want to visit a certain outdoor location your friend told you about. Your friend isn't available to take you there himself, but he has sent you *directions* for how to get to the place from the bus stop:

```
Sup G. Go to bus stop number 345. Bring a compass.
Walk 2 km north then 3 km east. You will find X there.
```

This text message contains all the information you need to find X.

Act 1: Following directions

You arrive at the bus stop, which is located at the top of a hill. From this height you can see the whole valley, and along the hillside below spreads a beautiful field of tall crops. The crops are so tall they prevent anyone standing in them from seeing too far; good thing you have a compass. You align the compass needle so the red arrow points north. You walk 2 km north, then turn 90° to the right

Vector as arrows

So far, we described how to perform algebraic operations on vectors in terms of their components. Vector operations can also be interpreted geometrically, as operations on arrows in the Cartesian plane.

Vector addition   The sum of two vectors corresponds to the combined displacement of the two vectors. Figure 3.4 illustrates the addition of two vectors, $\vec{v}_1 = (3, 0)$ and $\vec{v}_2 = (2, 2)$. The sum of the two vectors is the vector $\vec{v}_1 + \vec{v}_2 = (3, 0) + (2, 2) = (5, 2)$.
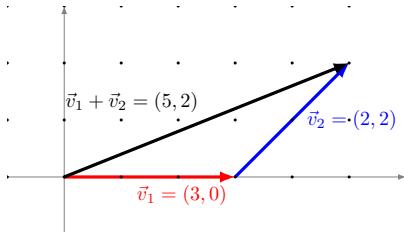
Figure 3.4: The addition of the vectors $\vec{v}_1$ and $\vec{v}_2$ produces the vector $(5, 2)$.

Vector subtraction   Before we describe vector subtraction, note that multiplying a vector by a scale factor $\alpha = -1$ gives a vector of the same length as the original, but pointing in the opposite direction.

This fact is useful if you want to subtract two vectors using the graphical approach. Subtracting a vector is the same as adding the negative of the vector:

$$\vec{w} - \vec{v}_1 = \vec{w} + (-\vec{v}_1) = \vec{v}_2.$$
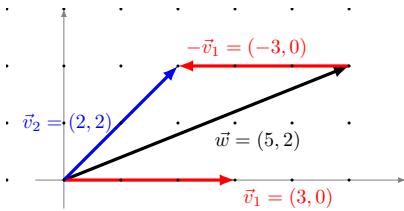
Figure 3.5: The vector subtraction $\vec{w} - \vec{v}_1$ is equivalent to the vector addition $\vec{w} + (-\vec{v}_1)$, where $(-\vec{v}_1)$ is like $\vec{v}_1$ but points in the opposite direction.

Figure 3.5 illustrates the graphical procedure for subtracting the vector $\vec{v}_1 = (3, 0)$ from the vector $\vec{w} = (5, 2)$. Subtraction of $\vec{v}_1 = (3, 0)$ is the same as addition of $-\vec{v}_1 = (-3, 0)$.

learning we often represent "rich data" like images, videos, and text as vectors with thousands of dimensions.

An example of an n-dimensional vector is

$$\vec{v} = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n.$$

The vector algebra operations you learned in this section also apply to these high-dimensional vectors.

Vectors and vector coordinates

One final point we need to clarify is the difference between real-world vector quantities like the velocity of a tennis ball $\vec{v}$ and its mathematical representation as a coordinate vector (v_x, v_y, v_z). If you know the coordinate vector (v_x, v_y, v_z) then you know what the real-world velocity is, right? Not quite.

Let's say you're doing a physics research project on tennis serves. You define an xyz-coordinate system for the tennis court, which allows you to represent the ball's velocity $\vec{v}$ as a triple of components (v_x, v_y, v_z) interpreted as: "The ball is moving with velocity v_x units in the x-direction, v_y units in the y-direction, and v_z units in the z-direction."

Suppose you want to describe the velocity vector $\vec{v}$ to a fellow physicist via text message. Referring to your sheet of calculations, you find the values $\vec{v} = (60, 3, -2)$, which you know were measured in metres per second. You send this message:

```
The velocity is (60,3,-2) measured in metres per second.
```

A few minutes later the following reply comes back:

```
Wait whaaat? What coordinate system are you using?
```

Indeed the information you sent is incomplete. Vector components depend on the coordinate system in which the vectors are represented. The triple of numbers $(60, 3, -2)$ only makes sense once you know the directions of the axes in the xyz-coordinate system. Realizing your mistake, you send a text with all the required information:

```
Using the coordinate system centred at the south post of
the net, with the x-axis pointing east along the court,
the y-axis pointing north along the net, and the z-axis
pointing up, the velocity is (60,3,-2) in metres per second.
```

A few seconds later, you get the reply:

```
OK got it now. Thx!
```

This hypothetical situation illustrates the importance of the coordinate systems for describing vectors. If you don't know what the coordinate system is, knowing the coordinates (v_x, v_y, v_z) doesn't tell you much. Only when you know the directions of the unit vectors $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$ can you interpret the instructions $\vec{v} = v_x\hat{\imath} + v_y\hat{\jmath} + v_z\hat{k}$.

It turns out, using the xyz-coordinate system with the three vectors $\{\hat{\imath}, \hat{\jmath}, \hat{k}\}$ is just one of many possible ways we can represent vectors. We can represent a vector $\vec{v}$ as coordinates (v_1, v_2, v_3) with respect to any *basis* $\{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$ using the expression $\vec{v} = v_1\hat{e}_1 + v_2\hat{e}_2 + v_3\hat{e}_3$, which corresponds to the instructions: "Move v_1 units in the direction of $\hat{e}_1$, move v_2 units in the direction of $\hat{e}_2$, and move v_3 units in the direction of $\hat{e}_3$."

What's a basis, you ask? I'm glad you asked, because this is the subject of the next section.

3.3 Basis

One of the most important concepts in the study of vectors is the concept of a *basis*. Consider the three-dimensional vector space $\mathbb{R}^3$. A *basis* for $\mathbb{R}^3$ is a set of vectors $\{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$ that can be used as a coordinate system for $\mathbb{R}^3$. If the set of vectors $\{\hat{e}_1, \hat{e}_2, \hat{e}_3\}$ is a basis, then you can *represent* any vector $\vec{v} \in \mathbb{R}^3$ as coordinates (v_1, v_2, v_3) *with respect to* that basis:

$$\vec{v} = v_1\hat{e}_1 + v_2\hat{e}_2 + v_3\hat{e}_3.$$

The vector $\vec{v}$ is obtained by measuring out a distance v_1 in the $\hat{e}_1$ direction, a distance v_2 in the $\hat{e}_2$ direction, and a distance v_3 in the $\hat{e}_3$ direction.

You are already familiar with the *standard* basis $\{\hat{\imath}, \hat{\jmath}, \hat{k}\}$, which is associated with the xyz-coordinate system. You know that any vector $\vec{v} \in \mathbb{R}^3$ can be expressed as a triple (v_x, v_y, v_z) with respect to the basis $\{\hat{\imath}, \hat{\jmath}, \hat{k}\}$ through the formula $\vec{v} = v_x\hat{\imath} + v_y\hat{\jmath} + v_z\hat{k}$. The whole point of this section is to let you know that other bases (coordinate systems) exist, and to get you into the habit of asking, "With respect to which coordinate system?" every time you see a coordinate vector (a, b, c).

An analogy

Let's start with a simple example of a basis. If you look at the HTML source code behind any web page, you're sure to find at least one mention of the colour stylesheet directive such as `color:#336699;`. The numbers should be interpreted as a triple of values $(33, 66, 99)$, each value describing the amount of red, green, and blue needed to

De Moivre's formula makes sense if you think of the complex number $z = e^{i\theta} = \cos\theta + i\sin\theta$, raised to the nth power:

$$(\cos\theta + i\sin\theta)^n = z^n = (e^{i\theta})^n = e^{in\theta} = \cos n\theta + i\sin n\theta.$$

Setting $n = 2$ in de Moivre's formula, we can derive the double angle formulas (page 98) as the real and imaginary parts of the following equation:

$$(\cos^2\theta - \sin^2\theta) + (2\sin\theta\cos\theta)i = \cos(2\theta) + \sin(2\theta)i.$$

Links

[Intuitive proof of the fundamental theorem of algebra]
https://www.youtube.com/watch?v=shEk8sz1oOw

3.6 Vectors problems

You learned a bunch of vector formulas and you saw some vector diagrams, but did you really learn how to solve problems with vectors? There is only one way to find out: test yourself by solving problems.

I've said it before and I don't want to repeat myself too much, but it's worth saying again: the more problems you solve, the better you'll understand the material. It's now time for you to try the following vector problems to make sure you're on top of things.

P3.1 Express the following vectors in length-and-direction notation:

    a) $\vec{u}_1 = (0, 5)$ **b)** $\vec{u}_2 = (1, 2)$ **c)** $\vec{u}_3 = (-1, -2)$

P3.2 Express the following vectors as components:

    a) $\vec{v}_1 = 20\angle 30°$ **b)** $\vec{v}_2 = 10\angle -90°$ **c)** $\vec{v}_3 = 5\angle 150°$

P3.3 Express the following vectors in terms of unit vectors $\hat{\imath}$, $\hat{\jmath}$, and $\hat{k}$:

    a) $\vec{w}_1 = 10\angle 25°$ **b)** $\vec{w}_2 = 7\angle -90°$ **c)** $\vec{w}_3 = (3, -2, 3)$

P3.4 Given the vectors $\vec{v}_1 = (1, 1)$, $\vec{v}_2 = (2, 3)$, and $\vec{v}_3 = 5\angle 30°$, calculate the following expressions:

    a) $\vec{v}_1 + \vec{v}_2$ **b)** $\vec{v}_2 - 2\vec{v}_1$ **c)** $\vec{v}_1 + \vec{v}_2 + \vec{v}_3$

P3.5 Starting from the point $P = (2, 6)$, the three displacement vectors shown in Figure 3.15 are applied to obtain the point Q. What are the coordinates of the point Q?

P3.6 Given the vectors $\vec{u} = (1, 1, 1)$, $\vec{v} = (2, 3, 1)$, and $\vec{w} = (-1, -1, 2)$, compute the following products:

    a) $\vec{u} \cdot \vec{v}$ **b)** $\vec{u} \cdot \vec{w}$ **c)** $\vec{v} \cdot \vec{w}$

    d) $\vec{u} \times \vec{v}$ **e)** $\vec{u} \times \vec{w}$ **f)** $\vec{v} \times \vec{w}$

Chapter 4

Mechanics

4.1 Introduction

Mechanics is the precise study of the motion of objects, the forces acting on them, and more abstract concepts such as momentum and energy. You probably have an intuitive understanding of these concepts already. In this chapter we will learn how to use precise mathematical equations to support your intuition.
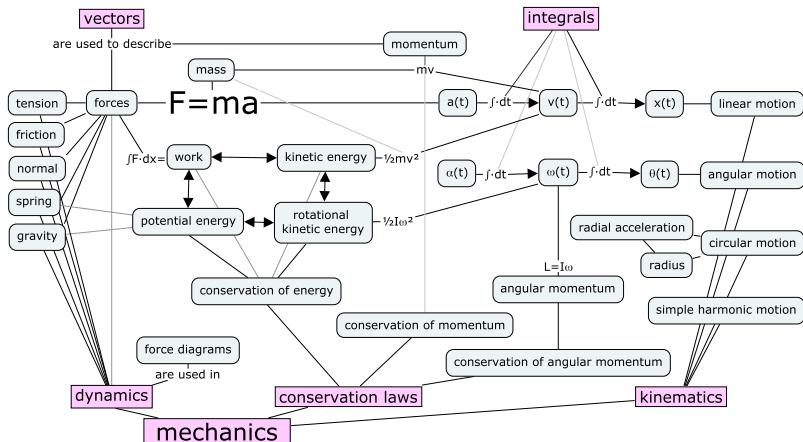

Figure 4.1: The concepts of mechanics. Forces are the cause of motion. We can also analyze the motion of objects in terms of the concepts of energy and momentum. If you understand the connections between all of the above concepts, you understand mechanics.

Newton's laws

Mechanics is the part of physics that is easiest to understand. Starting from three general principles known as *Newton's laws*, we can figure out pretty much everything about the motion of objects.

Newton's three laws of motion:

1. In the absence of external forces, objects will maintain their velocity and their direction of motion.

2. A force acting on an object causes an acceleration inversely proportional to the mass of the object: $\vec{F} = m\vec{a}$.

3. For each force $\vec{F}_{12}$ applied by Object 1 on Object 2, there is an equal and opposite force $\vec{F}_{21}$ that Object 2 exerts on Object 1.

The cool part of learning physics is that it teaches us to think about the laws of nature in terms of simple principles. Complicated phenomena can be broken down and understood in terms of basic theories. The laws of physics can be expressed in terms of mathematical equations. There are about 20 such equations (see page 510 in the back of the book). In this chapter you will learn how to use these equations to solve all kinds of physics problems.

Kinematics is the study of motion

To solve a physics problem is to obtain the *equation of motion* $x(t)$, which describes the position of the object as a function of time. Once you know $x(t)$, you can answer any question pertaining to the object's motion. To find the initial position x_i of the object, plug $t = 0$ into the equation of motion $x_i = x(0)$. To find the time(s) when the object reaches a certain distance, let's say 20[m] from the origin, solve for t in $x(t) = 20$[m]. Many of the problems on the mechanics final exam will be of this kind, so if you know how to find $x(t)$, you'll be in good shape to ace the exam.

In Chapter 2, we learned about the kinematics of objects moving in one dimension. More specifically, we used integration to obtain the velocity function of an object starting from the knowledge of its acceleration. Integrating the velocity function, we obtain its position function:

$$a(t) \quad \xrightarrow{v_i + \int dt} \quad v(t) \quad \xrightarrow{x_i + \int dt} \quad x(t).$$

Okay, but how do we obtain the acceleration?

Dynamics is the study of forces

The first step toward finding $x(t)$ is to calculate all the *forces* that act on the object. The forces are the *cause* of the object's acceleration. Newton's second law $F = ma$ states that **a force acting on an object produces an acceleration inversely proportional to the mass of the object**. There are many kinds of forces: the weight of an object $\vec{W}$ is a type of force, the force of friction $\vec{F}_f$ is another type of force, the tension in a rope $\vec{T}$ is yet another type of force, and so on. Note the little arrow on top of each force, which is there to remind you that forces are *vector quantities*. To find the *net force* acting on the object, calculate the sum of all the forces acting on the object $\vec{F}_\text{net} = \sum \vec{F}$.

Once you know the net force, you can use the formula $\vec{a}(t) = \frac{\vec{F}_\text{net}}{m}$ to find the object's acceleration. Once you know the acceleration $a(t)$, you can compute $x(t)$ using the calculus steps we learned in Chapter 2. The entire procedure for predicting the motion of objects can be summarized as

$$\frac{1}{m}\underbrace{\left(\sum \vec{F} = \vec{F}_\text{net}\right)}_{\text{dynamics}} = \underbrace{a(t) \xrightarrow{v_i + \int dt} v(t) \xrightarrow{x_i + \int dt} x(t)}_{\text{kinematics}}.$$

If you understand the above equation, then you understand mechanics. The goal of this chapter is to introduce you to all the concepts that appear in this equation and explore the relationships between them.

Other stuff

In addition to dynamics and kinematics, this chapter covers a number of other physics topics.

Newton's second law can also be applied to the study of objects in rotation. Angular motion is described by the angle of rotation $\theta(t)$, the angular velocity $\omega(t)$, and the angular acceleration $\alpha(t)$. Angular acceleration is caused by angular force, which we call *torque* $\mathcal{T}$. The principles behind circular motion are almost exactly the same as the principles of linear motion; the only difference being we use *angular* quantities to describe circular motion—instead of describing the motion in terms of [m], [m/s], and [m/s^2], we describe angular motion in terms of [radians], [radians/s], and [radians/s^2].

During a collision between two objects, the sudden spike in the contact force between them can be difficult to measure and quantify. It is therefore not possible to use Newton's law $F = ma$ to find the accelerations of the objects occurring during collisions and predict the motion of the objects using the kinematics approach described above.

To predict the motion of objects after a collision, we can use a *momentum* calculation. An object of mass m moving with velocity $\vec{v}$ has momentum $\vec{p} = m\vec{v}$. The principle of conservation of momentum states that **the total amount of momentum in a system before and after a collision is conserved**. Thus, if two objects with initial momenta $\vec{p}_{i1}$ and $\vec{p}_{i2}$ collide, the total momentum before the collision must be equal to the total momentum after the collision:

$$\sum \vec{p}_i = \sum \vec{p}_f \qquad \Rightarrow \qquad \vec{p}_{i1} + \vec{p}_{i2} = \vec{p}_{f1} + \vec{p}_{f2}.$$

We use this equation to calculate the final momenta $\vec{p}_{f1}$, $\vec{p}_{f2}$ of the objects after the collision.

There is another way to solve physics problems by applying the concept of *energy*. Instead of trying to describe the entire motion of the object, we can focus only on the initial parameters and the final parameters of an object's movement. The law of conservation of energy states that **the total energy of the system is conserved**:

$$\sum E_i = \sum E_f.$$

By knowing the total initial energy of a system, we can find the final energy in the system, and from the final energy we can calculate the final motion parameters.

Units

In math we work with numbers—we solve questions where the answers are numbers without dimensions like 3, 5 or 12.34. The universal power of math comes precisely from this abstraction of things into numbers. We could be solving for the number of sheep in a pen, the surface area of a sphere, or the annual revenue of your startup; we can apply the same mathematical techniques to each example, even though the numbers we use will represent very different kinds of quantities.

Because physics deals with real-world concepts and quantities, each number in physics always comes with a *measurement unit*. We must pay attention to the units of physical quantities and—most importantly—distinguish between the different *dimensions* of numerical quantities. An answer in physics is a number that represents a length, a time, a velocity, an acceleration, or some other physical quantity. It doesn't make sense to add a *time* and a *mass*, because the two numbers measure different kinds of quantities.

Here's a list of some kinds of quantities discussed in this chapter:

Dimension	SI unit	Other units	Measured with
time	[s]	[h], [min]	clock
length	[m]	[cm], [mm], [ft], [in]	metre tape
velocity	[m/s]	[km/h], [mi/h]	speedometer
acceleration	[m/s^2]		accelerometer
mass	[kg]	[g], [lb]	scale

Appendix C (see page 478 in the back of the book) provides a more detailed list of the International System of Units (abbreviated SI for *Système International*).

The units of physical quantities are indicated in square brackets throughout the lessons of this chapter. In your equations, you should always try to keep in mind the units for different physics quantities. Sometimes you'll be able to catch yourself making an error because the units will not come out right. If I ask you to calculate the maximum height a ball will reach, I expect your answer to be a length measured in [m] and not some other kind of quantity like a velocity [m/s] or an acceleration [m/s^2] or an area [m^2]. An answer in [ft] would also be acceptable since this is also a length, and it can be converted to metres using $1[ft] = 0.3048[m]$ (see page 479 for other conversion ratios). Learn to watch out for the units and dimensions of physical quantities, and you'll have an easy time in physics. They are an excellent error-checking mechanism.

We'll begin our physics journey by starting with the familiar subject of kinematics which we studied in Chapter 2. Now that you know about vectors, we can study two-dimensional kinematics problems, such as the motion of a projectile.

4.2   Projectile motion

Ever since the invention of gun powder, generation after generation of men have thought of countless different ways to hurtle shrapnel at each other. Indeed, mankind has been stuck to the idea of two-dimensional projectile motion like flies on shit. As long as there is money to be made in selling weapons, and so long as the media continues to justify the legitimacy of the use of these weapons, it is likely the trend will continue.

It is therefore imperative for anyone interested in reversing this trend to learn about the physics of projectile motion. You need to know the techniques of the enemy (the industrial military complex) before you can fight them. We'll see that projectile motion is nothing
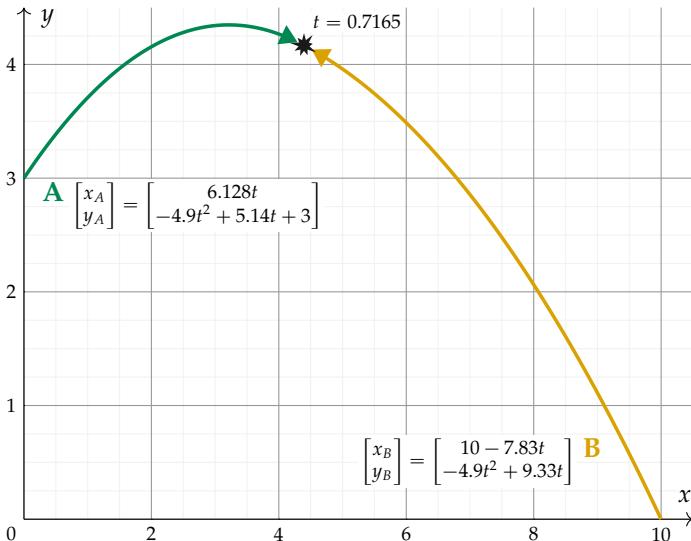
Figure 4.3: The incoming ball (A) is intercepted by the ball (B) coming from the right. The collision occurs at $t = 0.7165$[s].

We express these conditions through the following two equations:

$$8\cos(40)t + 0 = -w\cos(50)t + 10,$$
$$-4.9t^2 + 8\sin(40)t + 3 = -4.9t^2 + w\sin(50)t + 0,$$

which we must solve simultaneously.

To solve, we cancel $-4.9t^2$ on both sides of the bottom equation:

$$8\cos(40)t = -w\cos(50)t + 10,$$
$$8\sin(40)t + 3 = w\sin(50)t.$$

This is a set of two equations with two unknowns, and we can solve it. It's not going to be easy, since we can't cleanly isolate t or w using standard substitution techniques. However, there is a trick! We can divide the two equations. If $A = B$ and $C = D \neq 0$ then $A/C = B/D$, and this is what we'll use. To prepare for this step, let's rearrange the equations a bit so all the w-containing terms stand alone on the right side:

$$10 - 8\cos(40)t = w\cos(50)t,$$
$$8\sin(40)t + 3 = w\sin(50)t.$$

We'll now divide the bottom equation by the top equation to obtain

$$\frac{8\sin(40)t + 3}{10 - 8\cos(40)t} = \frac{w\sin(50)t}{w\cos(50)t} = \tan(50).$$
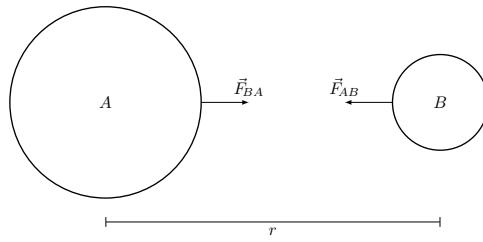
Figure 4.4: The gravitational force between two planets acts to pull them to-gether. Assume Planet A has mass m_A and Planet B has mass m_B. The force vector $\vec{F}_{AB}$ describes "Planet A's pull on Planet B." The force vector $\vec{F}_{BA}$ describes "Planet B's pull on Planet A." The magnitude of the gravitational pull is $F_g = \frac{Gm_Am_B}{r^2} = \|\vec{F}_{AB}\| = \|\vec{F}_{BA}\|$.

mass m is given by

$$F_g = \frac{GMm}{r^2} = \underbrace{\frac{GM}{r^2}}_{g} m = 9.81m = W.$$

We call this force the *weight* of the object, and to be precise we write $\vec{W} = -mg\hat{\jmath}$ to indicate that the weight acts *downward*, in the negative y-direction. Verify using a calculator that $\frac{GM}{r^2} = 9.81 = g$.

Force of a spring

A spring is a piece of metal twisted into a coil that has a certain natu-ral length. The spring will resist any attempts to stretch or compress it. The force exerted by a spring is given by

$$\vec{F}_s = -kx,$$

where x is the amount by which the spring is displaced from its nat-ural length, and the constant k[N/m] is a measure of the spring's *strength*. Note the negative sign indicates that the spring always acts to oppose the displacement.

If you try to stretch the spring, displacing it in the positive x-direction, then the force of the spring will pull against you (the spring will pull in the negative x-direction). Similarly, if you try to compress the spring (a displacement in the negative x-direction), the spring will push back against you, in the positive x-direction.

Normal force

The *normal* force is the force between two surfaces in contact. In this context, the word *normal* means "perpendicular to the surface of."
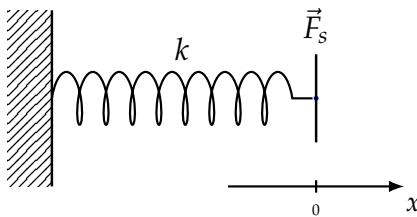
Figure 4.5: The force exerted by a spring is proportional to its displacement from its natural length, denoted x. When $x > 0$ the spring is stretched. When $x < 0$ the spring is compressed. When one end of the spring is fixed, the force exerted by the spring at the other end is $\vec{F}_s = -kx$.

The reason my coffee mug is not falling to the floor right now is that the table exerts a normal force $\vec{N}$ on the mug, keeping it in place.

Force of friction

In addition to the normal force between surfaces, there is also the force of friction $\vec{F}_f$, which acts to impede any sliding motion between the surfaces. There are two kinds of friction forces, and both are proportional to the amount of normal force between the surfaces:

$$\max\{\vec{F}_{fs}\} = \mu_s \|\vec{N}\| \text{ (static)}, \quad \text{and} \quad \vec{F}_{fk} = \mu_k \|\vec{N}\| \text{ (kinetic)},$$

where μ_s and μ_k are the static and kinetic *friction coefficients*. It makes sense that the force of friction should be proportional to the magnitude of the normal force $\|\vec{N}\|$, since the harder the two surfaces push against each other, the more difficult it becomes to make them slide. The above equations give mathematical precision to this intuitive logic.

The static force of friction acts on objects that are not moving. It describes the *maximum* amount of friction that can exist between two objects. If a horizontal force greater than $F_{fs} = \mu_s N$ is applied to the object, then it will start to slip. The kinetic force of friction acts when two objects are sliding relative to each other. It always acts in the direction opposite to the motion.

Tension

A force can also be exerted on an object remotely by attaching a rope to the object, and pulling the rope. The force exerted on the object will be equal to the rope's *tension* $\vec{T}$. Note that tension always pulls *away* from an object: you can pull but you can't push a dog by its leash.
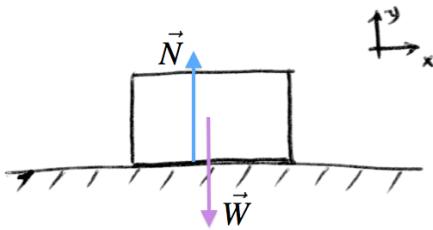
Figure 4.6: A block sitting on a table. The weight of the block $\vec{W}$ is counteracted by the normal force $\vec{N}$.

Step 3: Next, write the empty force diagram equations template:

$$\sum F_x = \qquad = ma_x,$$
$$\sum F_y = \qquad = ma_y.$$

Step 4: There are no forces acting in the x-direction, and the block is not moving, so $a_x = 0$. In the y-direction, we have the force of gravity and the normal force exerted by the table:

$$\sum F_x = 0 = 0,$$
$$\sum F_y = N - mg = 0.$$

We set $a_y = 0$, as we can *see* that the block is just sitting there on the table without moving. The technical term for situations where $a_x = 0, a_y = 0$ is called *static equilibrium*. Force diagrams with static equilibrium are easy to solve because the entire right-hand side is equal to zero, which means the forces acting on the object must be counter-balancing each other.

Step 5: Suppose the teacher now asks, "What is the magnitude of the normal force?" By looking at the second equation, you can answer, "$N = mg$ bro!"

Moving the fridge

You are trying to push your fridge across the kitchen floor. It weighs quite a lot, and is strongly "gripping" the floor when you try to push it. The static coefficient of friction between the bottom of your fridge and the tiles on the floor is μ_s. How much force $\vec{F}_{\text{ext}}$ will it take to cause the fridge to start moving?

Step 5: We know the force exerted by a spring is proportional to its displacement according to

$$F_s = -ky_B,$$

so we can find $y_B = -\frac{mg}{k}$. The length of compression is therefore

$$|\Delta y| = \frac{mg}{k}.$$

Two blocks

Now you're ready for a more involved example with two blocks. One block is sitting on a surface, and the other block is falling straight down. The two blocks are connected by a rope. What is the acceleration of the *system* as a whole?

Steps 1 and 2: We have two objects, so we need to draw two force diagrams.
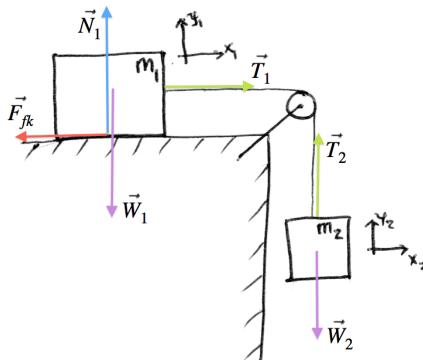
Figure 4.11: A block of mass m_1 is dragged along horizontally by a second block of mass m_2 which is falling vertically. What is the acceleration of the system?

Step 3: We need two sets of equations; one set for the block on the horizontal surface, and one set for the falling block:

$$\sum F_{1x} = \qquad = m_1 a_{x_1} \qquad\qquad \sum F_{2x} = \qquad = m_2 a_{x_2}$$
$$\sum F_{1y} = \qquad = m_1 a_{y_1} \qquad\qquad \sum F_{2y} = \qquad = m_2 a_{y_2}$$

Step 4: We fill in the equations with all the forces drawn in the diagram:

$$\sum F_{1x} = -F_{fk} + T_1 = m_1 a_{x_1} \qquad\qquad \sum F_{2x} = 0 = 0$$
$$\sum F_{1y} = N_1 - W_1 = 0 \qquad\qquad \sum F_{2y} = -W_2 + T_2 = m_2 a_{y_2}$$

We can summarize the entire procedure for predicting the position of an object $x(t)$ from first principles in the following equation:

$$\frac{1}{m} \underbrace{\left(\sum \vec{F} = \vec{F}_{\text{net}} \right)}_{\text{dynamics}} = \underbrace{\vec{a}(t) \xrightarrow{\vec{v}_i + \int dt} \vec{v}(t) \xrightarrow{\vec{r}_i + \int dt} \vec{r}(t)}_{\text{kinematics}} .$$

The left-hand side calculates the net force acting on an object, which is the *cause* of acceleration. The right-hand side indicates how we can calculate the position vector $\vec{r}(t)$ starting from the acceleration and the initial conditions. If you know the forces acting on any object (rocks, projectiles, cars, stars, planets, etc.) then you can predict the object's motion using this equation, which is pretty cool.

So far we discussed one approach for analyzing the motion of objects. Calculating the forces and the acceleration of objects, then using integration to find the position function $\vec{r}(t)$ is a very useful approach for solving physics problems. There are several other ways of looking at the motion of objects that are equally useful and provide us with different insights. In the next two sections, we'll discuss how to model physical situations in terms of momentum and energy.

4.5   Momentum

A collision between two objects creates a sudden spike in the contact force between them, which can be difficult to measure and quantify. It is not possible to use Newton's law $F = ma$ to predict the accelerations that occur during collisions. To predict the motion of the objects after the collision, we need a *momentum* calculation. According to the law of conservation of momentum, the total amount of momentum before and after the collision is the same. Once we know the momenta of the objects before the collision, it becomes possible to calculate their momenta after the collision, and from this determine their subsequent motion.

To illustrate the importance of momentum, consider the following situation. Say you have a 1[g] paper ball and a 1000[kg] car moving at the same speed of 100[km/h]. Which of the two objects would you rather be hit by? Momentum, denoted $\vec{p}$, is the precise physical concept that measures the *quantity* of motion. An object of mass m moving with velocity $\vec{v}$ has a momentum of $\vec{p} = m\vec{v}$. Momentum plays a key role in collisions. Your gut feeling about the piece of paper and the car is correct. The car weighs $1000 \times 1000 = 10^6$ times more than the piece of paper, so the car has 10^6 times more momentum when moving at the same speed. Colliding with the car
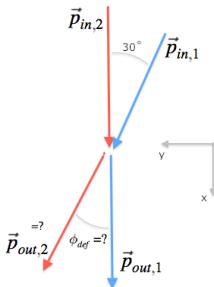
Figure 4.12: Momentum diagram of the collision between two cyclists.

Assume the weight of the street-crossing-at-30-degrees hipster combined with the weight of his bike totals 90[kg]. And assume the weight of the hipster moving in the straight line combined with the weight of his bike is 100[kg].

The story will continue in a moment, but first let's review the information I've given you so far:

$$\vec{p}_{in,1} = 90 \times 50\angle 30$$
$$= 90(50\cos 30, 50\sin 30),$$

$$\vec{p}_{in,2} = 100 \times 50\angle 0$$
$$= 100(50, 0),$$

where the x-coordinate points down the street, and the y-coordinate is perpendicular to the street.

Surprisingly, nobody gets hurt in this collision. The bikers bump shoulder-to-shoulder and bounce off each other. The hipster who was trying to cross the street is redirected down the street, while the hipster travelling down the street is deflected to the side and rerouted onto a bike path. I know what you are thinking: couldn't they get hurt at least a little bit? Okay, let's say the whiplash from the shoulder-to-shoulder collision sends the hipsters' heads flying toward each other and smashes their glasses. There you have it.

Suppose the velocity of the first hipster after the collision is 60 [km/h]. What is the velocity and the deflected direction of the second hipster? As given above, the outgoing momentum of the first hipster is $\vec{p}_{out,1} = 90(60, 0)$, and we're looking to find $\vec{p}_{out,2}$.

We can solve this problem with the conservation of momentum formula, which tells us that

$$\vec{p}_{in,1} + \vec{p}_{in,2} = \vec{p}_{out,1} + \vec{p}_{out,2}.$$

again. Complicated as these forces may be, we know that during the entire collision they obey Newton's third law. Assuming there are no other forces acting on the objects, we have

$$\vec{F}_{12} = -\vec{F}_{21} \quad \text{using the above} \Rightarrow \quad \frac{d\vec{p}_1}{dt} = -\frac{d\vec{p}_2}{dt}.$$

If we move the negative term to the left-hand side of the equation we obtain

$$\frac{d\vec{p}_1}{dt} + \frac{d\vec{p}_2}{dt} = 0 = \frac{d}{dt}\left(\vec{p}_1 + \vec{p}_2\right).$$

The second part of the equation implies that the quantity $(\vec{p}_1 + \vec{p}_2)$ is constant over time, and so $\vec{p}_{\text{in},1} + \vec{p}_{\text{in},2} = \vec{p}_{\text{out},1} + \vec{p}_{\text{out},2}$.

In this section, we saw how to use a momentum calculation to predict the motion of particles after a collision. In the next section we'll learn about *energy*, which is another useful concept for understanding and predicting the motion of objects.

Links

[Animations of simple collisions between objects]
`http://en.wikipedia.org/wiki/Conservation_of_linear_momentum`

Exercise

E4.1 A sticky ball of mass 3[g] and velocity 20[m/s] collides with a stationary ball of mass 5[g]. The balls stick together. What is their velocity after the collision?
Hint: Use conservation of momentum $\vec{p}_{1,\text{in}} + \vec{p}_{2,\text{in}} = \vec{p}_{\text{out}}$.

4.6 Energy

Instead of thinking in terms of velocities $v(t)$ and motion trajectories $x(t)$, we can solve physics problems by using *energy* calculations. In this section, we'll precisely define different kinds of energies, and we'll learn the rules for converting one energy into another. The key idea to keep in mind is the principle of *total energy conservation*, which says that in any physical process, the sum of the initial energies is equal to the sum of the final energies.

The negative of the work done against a conservative force is called *potential energy*. For any conservative force $\vec{F}_?$, we can define the associated potential energy $U_?$ through the formula,

$$U_?(d) = -W_{\text{done}} = -\int_0^d \vec{F}_? \cdot d\vec{x}.$$

We'll discuss two specific examples of this general formula below: gravitational potential energy and spring potential energy. An object high in the air has a great potential to fall; similarly, compressing a spring by a certain distance gives it the potential to spring back to its normal position. Let's look at the exact formulas for these two cases.

Gravitational potential energy

The force of gravity is given by $\vec{F}_g = -mg\hat{\jmath}$. The direction of the gravitational force is downward, toward the centre of the Earth.

The gravitational potential energy of lifting an object from a height of $y = 0$ to a height of $y = h$ is given by

$$U_g(h) = -W_{\text{done}} = -\int_0^h \vec{F}_g \cdot d\vec{y} = -\int_0^h (-mg\hat{\jmath}) \cdot \hat{\jmath} \, dy$$

$$= mg \int_0^h 1 \, dy = mgy \Big|_{y=0}^{y=h}$$

$$= mgh \qquad [\text{J}].$$

Spring potential energy

The force of a spring when displaced by $\vec{x}$[m] from its natural position is given by $\vec{F}_s(\vec{x}) = -k\vec{x}$. The potential energy stored in a spring as it is stretched from $y = 0$ to $y = x$[m] is given by

$$U_s(x) = -W_{\text{done}} = -\int_0^x \vec{F}_s(y) \cdot d\vec{y}$$

$$= -\int_0^x (-ky) \, dy = k \int_0^x y \, dy = k\frac{1}{2}y^2 \Big|_{y=0}^{y=x}$$

$$= \frac{1}{2}kx^2 \qquad [\text{J}].$$

Note the formula applies when $x > 0$ (spring stretched by length $|x|$) and when $x < 0$ (spring compressed by length $|x|$).

4.8 Angular motion

We'll now study the physics of objects in rotation. Rotational motion is exemplified by spinning disks, rotating bicycle wheels, spinning footballs, and spinning figure skaters, among other spinning things.

As you'll see shortly, the basic concepts we'll use to describe angular motion are directly analogous to the concepts of linear motion: position, velocity, acceleration, force, momentum, and energy.

Review of linear motion

It will be helpful to begin with a quick review of the concepts and formulas used to describe the linear motion of objects.

The linear motion of an object is described by its position $x(t)$, velocity $v(t)$, and acceleration $a(t)$ as functions of time. The position function tells you where the object is, the velocity tells you how fast it is moving, and the acceleration measures the change in the object's velocity.

The motion of objects is governed by Newton's first and second laws. In the absence of external forces, objects will maintain a uniform velocity (UVM), which corresponds to the equations of motion $x(t) = x_i + v_i t$ and $v(t) = v_i$. If a net force $\vec{F}$ acts on the object, the force will cause the object to accelerate. We obtain the magnitude of this acceleration with the formula $F = ma$. A constant force acting on an object will produce a constant acceleration (UAM), which corresponds to the equations of motion $x(t) = x_i + v_i t + \frac{1}{2}at^2$ and $v(t) = v_i + at$.
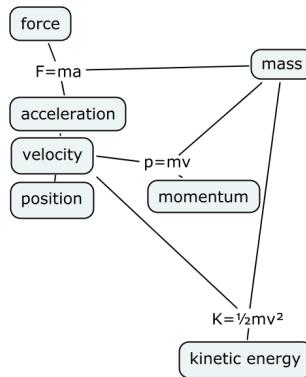

Figure 4.18: The main concepts used to describe linear motion.

We also learned how to quantify the *momentum* $\vec{p} = m\vec{v}$ and the
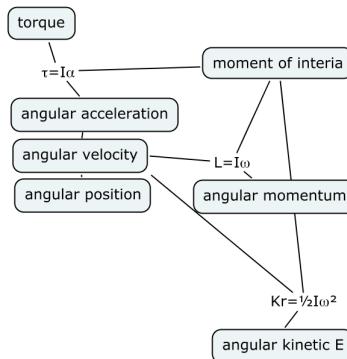
Figure 4.21: The main concepts used to describe angular motion.

inertia also appears in the formulas for finding a spinning object's angular momentum and its rotational kinetic energy.

4.9 Simple harmonic motion

Vibrations, oscillations, and waves are everywhere around us. For example, what appears to our eyes as white light is actually made of many different oscillations of the electromagnetic field. These oscillations vibrate at a range of frequencies, which correspond to the colours we perceive. Sounds are also made of combined air vibrations with various frequencies and strengths. In this section, we'll learn about *simple harmonic motion*, which describes the oscillation of a mechanical system at a fixed frequency and with a constant amplitude. As its name suggests, simple harmonic motion is the simplest form of oscillatory motion. By studying oscillations in their simplest form, you'll gather important intuition that applies to all types of oscillations and wave phenomena.
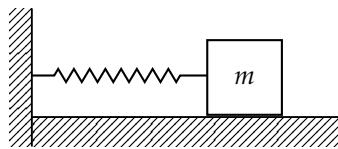

Figure 4.22: A block of mass m attached to a spring.

The canonical example of simple harmonic motion is the motion of a mass-spring system, as illustrated in Figure 4.22. The block is free to slide along the horizontal frictionless surface. If the system is

Our goal was to find $x(t)$ for all times t. However, none of the equations of physics tell us $x(t)$ directly. Instead, we have Newton's second law $F = ma$, which tells us that the acceleration of the object $a(t)$ equals the *net force* acting on the object divided by the object's mass. To find $x(t)$ starting from $a(t)$, we use integration twice:

$$\frac{1}{m}\left(\sum \vec{F} = \vec{F}_{\text{net}}\right) = a(t) \xrightarrow{v_i + \int dt} v(t) \xrightarrow{x_i + \int dt} x(t).$$
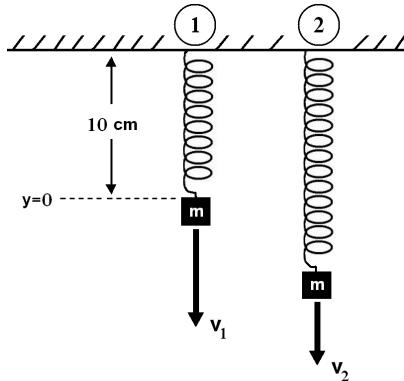
We studied kinematics in several different contexts. We originally looked at kinematics problems in one dimension, and derived the UAM and UVM equations. We also studied the problem of projectile motion by deconstructing it into two separate kinematics subproblems: one in the x-direction (UVM), and one in the y-direction (UAM). Later, we studied the circular motion of objects and stated equation $a_r = \frac{v_t^2}{r}$, which describes an important relationship between the radial acceleration, the tangential velocity, and the radius of the circle of rotation. We also studied rotational motion using angular kinematics quantities $\theta(t)$, $\omega(t)$, and $\alpha(t)$. We defined the concept of *torque* and saw the role it plays in the angular equivalent of Newton's second law $\mathcal{T} = I\alpha$. We studied the equation that describes simple harmonic motion, $x(t) = A\cos(\omega t + \phi)$, and showed the formula $\omega = \sqrt{\frac{k}{m}}$, which gives the angular frequency of a mass-spring system.

We also discussed three conservation laws: the conservation of linear momentum law $\sum \vec{p}_i = \sum \vec{p}_f$, the conservation of angular momentum law $L_i = L_f$, and the conservation of energy law $\sum E_i = \sum E_f$. Each of these three fundamental quantities is conserved overall and can neither be created nor destroyed. Momentum calculations are used to analyze collisions, while energy formulas like equations $K = \frac{1}{2}mv^2$, $U_g = mgh$, and $U_s = \frac{1}{2}kx^2$ can be used to analyze the motion of objects in terms of energy principles.

Now you can see how 20 equations truly are enough to master all of mechanics. Nice work! Your next step should be to practice solving some problems in order to solidify your understanding.

4.11 Mechanics problems

It's now time for you to verify experimentally how well you've understood the material from this chapter. Try solving the physics problems presented in this section. Go ahead, dig in! And don't be discouraged if you find some of the problems difficult—they are meant to be challenging in order to force you to think hard and reinforce the connections between the concepts in your head.
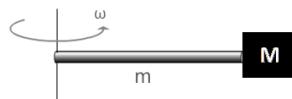
P4.6 You throw a ball from the ground vertically at speed v and measure its speed when it falls back to the ground. You first carry out this experiment on Earth, then repeat it on the Moon. Does the ball have a greater speed as it hits the ground on Earth or on the Moon?
Hint: Use conservation of energy.

P4.7 In the previous problem, assume there is a pit that allows the ball to fall 10[m] below the level from which it was thrown. Will the ball have a greater speed on Earth or on the Moon when it hits the bottom of the 10[m]-deep pit?
Hint: Use conservation of energy.

P4.8 A rod of mass m is rotating horizontally about one of its ends. An additional mass M is attached at the other end. Assume the system rotates at a constant angular velocity ω. What is the torque on the mass M? If the mass M is detached from the rod without any intervention of an external force, what will the new angular velocity of the rod be?

Hint: Use $\mathcal{T} = I\alpha$. Use conservation of angular momentum.

P4.9 A car is driving inside a vertical circular loop-de-loop. Assume the car passes the top and the bottom of the loop at the same speed. Will the normal force exerted by the loop on the car be greater at the top or at the bottom?
Hint: The car requires a centripetal force to maintain its circular path.

P4.10 Two balls of mass m are thrown from the top of a building with equal velocity $\vec{v}$ as shown in the diagram. Which ball has the greatest speed at the moment it hits the ground (ignore air resistance)?

1. Find the friction coefficient μ_s in terms of m_1, m_2, and M.

2. Calculate the friction coefficient μ_s if $m_1 = 3.50[\text{kg}]$, $m_2 = 6.00[\text{kg}]$, and $M = 4.00[\text{kg}]$.

3. If there were no friction ($\mu_s = 0$, $\mu_k = 0$), what would the acceleration of mass the M be, in terms of m_1, m_2, M, and g?
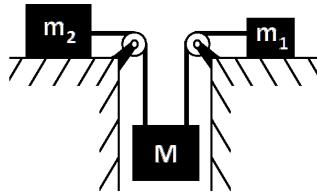

Figure 4.30: The three-mass system analyzed in **P4.25**.

Hint: Draw a force diagram for each block. The blocks have the same a.

P4.26   A ball is fired with $v_{1i} = 10[\text{m/s}]$ at an angle $\theta = 30°$ with the horizontal. You need to fire a second ball vertically from the same height such that it hits the first ball as it reaches its maximum height.
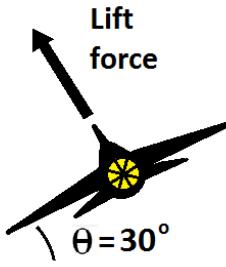
1. Find the horizontal distance d from the first ball's firing position, from where you should fire the second ball.

2. Find the initial velocity required for the second ball.

Hint: Find the maximum height and the half-range of the first ball.

P4.27   A $14\,000[\text{kg}]$ F-16 fighter jet is in a dogfight. The pilot needs to make a turn with a radius of $5[\text{km}]$ while maintaining a speed of $605.5[\text{km/h}]$. The plane executes a banked turn at an angle $\theta = 30°$ and follows a horizontal circular path.

1. How much lift does the pilot need to perform this maneuver? The lift force is perpendicular to the wings and the fuselage of the plane.

2. Will the altitude of the fighter change during this maneuver?

Hint: Find the lift force needed to produce the centripetal acceleration.

P4.28   You're in a subway car moving at $v = 12.5[m/s]$ when you drop a water bottle on the floor. The bottle comes to rest with respect to the subway car. The subway then starts braking and comes to a stop. The bottle starts rolling forward without slipping on the floor of the subway car. Find the linear velocity of the bottle as it rolls forward. The bottle's moment of inertia is $I = \frac{1}{2}mr^2$ and its mass is m.
Hint: The kinetic energy of the bottle is conserved.

P4.29   You're playing with two hockey pucks on a pool table as shown in Figure 4.31. The coefficient of friction between the pucks and the table is μ_k. Puck 2 is at rest before the collision and at a distance d from the corner pocket. Puck 1 hits Puck 2 ($m_1 = m_2$) with velocity v_i. After the collision, Puck 1 has velocity v_1 and Puck 2 has velocity v_2.

1. What is the minimum v_i in terms of the variables provided such that Puck 2 enters the pocket?

2. Calculate v_i, v_1, and v_2 if $\mu_k = 0.273$ and $d = 0.70[m]$.

3. How far does Puck 1 move after the collision if v_i is chosen as in part 2 of the question?

4. Is the collision elastic?

Hint: Use conservation of momentum and $K_i = K_f + W_{lost}$.
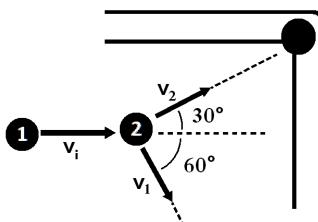
Figure 4.31: The collision of two pucks analyzed in **P4.29**.

P4.30   A car accelerates at $a = 4[m/s^2]$. Each tire has radius $r = 30[cm]$ and moment of inertia $I = 0.27[kg\,m^2]$. The car doesn't skid.

1. What is the torque applied on each tire?

2. If $\theta_i = 0$ and $\omega_i = 3[rad/s]$, how many revolutions did the tire complete from $t = 0[s]$ until $t = 4[s]$?

Hint: Use the relation between a and α.

P4.31   A solid cylinder and a hollow cylinder of identical mass are placed side by side on an incline. If both cylinders are released from rest and start rolling, which cylinder will reach the bottom of the incline first?
Hint: Think about $\mathcal{T} = I\alpha$.

P4.32   A ball is thrown from ground level with upward initial velocity $20[m/s]$. How long will the ball be in the air before it returns to the ground?

Chapter 5

Calculus

Calculus is *useful* math. We use calculus to solve problems in physics, chemistry, computing, biology, and many other areas of science. You need calculus to perform the quantitative analysis of how functions change over time (derivatives), and to calculate the total amount of a quantity that accumulates over a time period (integrals).

The language of calculus will allow you to speak precisely about the properties of functions and better understand their behaviour. You will learn how to calculate the slopes of functions, how to find their maximum and minimum values, how to compute their integrals, and other tasks of practical importance.

5.1 Introduction

In Chapter 2, we developed an intuitive understanding of integrals. Starting with the knowledge of an object's acceleration function over time, we used the integration operation to calculate the object's velocity function and its position function. We'll now take a closer look at the techniques of calculus using precise mathematical statements, and study how these techniques apply to other problems in science.

A strong knowledge of functions is essential for your understanding of the new calculus concepts. I recommend revisiting Section 1.12 (page 63) to remind yourself of the functions introduced therein. I insist on this. Go! Seriously, there is no point in learning that the derivative of the function $\sin(x)$ is the function $\cos(x)$ if you don't have a clue what $\sin(x)$ and $\cos(x)$ are.

Before we introduce any formal definitions and formulas, let's demonstrate how calculus is used in a real-world example.

Download example

Suppose you're downloading a large file to your computer. At $t = 0$ you click "save as" in your browser and the download starts. Let $f(t)$ represent the size of the downloaded data. At any time t, the function $f(t)$ tells you the amount of disk space taken by the partially-downloaded file. You are downloading a 720[MB] file, so the download progress at time t corresponds to the fraction $\frac{f(t)}{720[\text{MB}]}$.

Download rate

The derivative function $f'(t)$, pronounced "f prime," describes how the function $f(t)$ changes over time. In our example $f'(t)$ is the download speed. If your downloading speed is $f'(t) = 100[\text{kB/s}]$, then the file size $f(t)$ must increase by 100[kB] each second. If you maintain this download speed, the file size will grow at a constant rate: $f(0) = 0[\text{kB}]$, $f(1) = 100[\text{kB}]$, $f(2) = 200[\text{kB}]$, ..., $f(100) = 10[\text{MB}]$.

To calculate the "estimated time remaining" until the download's completion, we divide the amount of data that remains to be downloaded by the current download speed:

$$\text{time remaining} \; = \; \frac{720 - f(t)}{f'(t)} \quad [\text{s}].$$

The bigger the derivative, the faster the download will finish. If your internet connection were 10 times faster, the download would finish 10 times more quickly.

Inverse problem

Let's consider this situation from the point of view of the modem that connects your computer to the internet. Any data you download comes through the modem. The modem knows the download rate $f'(t)[\text{kB/s}]$ at all times during the download.

However, since the modem is separate from your computer, it does not know the file size $f(t)$ as the download progresses. Nevertheless, the modem can infer the file size at time t from knowing the transmission rate $f'(t)$. The integral of the download rate between $t = 0$ and $t = \tau$ corresponds to the total amount of downloaded data stored on your computer. During this download period, the change in file size is described by the integral

$$\Delta f = f(\tau) - f(0) = \int_0^\tau f'(t)\, dt\,.$$

Assuming the file size starts from zero $f(0) = 0[\text{kB}]$ at $t = 0$, the modem can use the integration procedure to find $f(\tau)$, the file size on your computer at $t = \tau$:

$$f(\tau) = \int_0^\tau f'(t)dt.$$

The download rate $f'(t)$ is measured in $[\text{kB/s}]$, and each time step dt is 1[s] long, so the data downloaded during one second is $f'(t)dt[\text{kB}]$. The file size at time $t = \tau$ is equal to the sum of the data downloaded during each second from $t = 0$ until $t = \tau$.

The integral $\int_a^b q(t)\,dt$ is the calculation of the *total* of some quantity $q(t)$ that accumulates during the time period from $t = a$ to $t = b$. Integrals are necessary any time you want to calculate the total of a quantity that changes over time.

As demonstrated above, calculus is much more than the theoretical activity reserved for math specialists. Calculus relates to everyday notions you're already familiar with. Indeed, we carry out calculus-like operations in our head every day—we just don't necessarily use calculus terminology when we do so.

Learning the language of calculus will help you think more clearly about certain types of problems. Understanding the language of calculus is *essential* for learning science because many laws of nature are best described in terms of derivatives and integrals.

Usually, differential calculus and integral calculus are taught as two separate subjects. Perhaps teachers and university administrators are worried the undergraduates' little heads will explode from sudden exposure to *all* of calculus. However, this separation actually makes calculus more difficult, and prevents students from discovering the connections between differential and integral calculus. We'll have no such split in this book, because I believe you can handle the material in one go. Understanding calculus involves figuring out new mathematical concepts like infinity, limits, and summations, but these ideas are not *that* complicated. By getting this far, you've proven you're more than ready to learn the theory, techniques, and applications of derivatives, integrals, sequences, and series.

Let's begin with an overview of the material.

5.2   Overview

This section presents a bird's-eye view of the core concepts of calculus. We'll define more precisely the operations of differentiation and integration, which were introduced in Chapter 2 (see page 165).

We'll also discuss the other parts of calculus: *limits*, *sequences*, and *series*. We'll briefly touch upon some applications for each of these concepts; after all, you should know *why* you want to learn all this stuff.
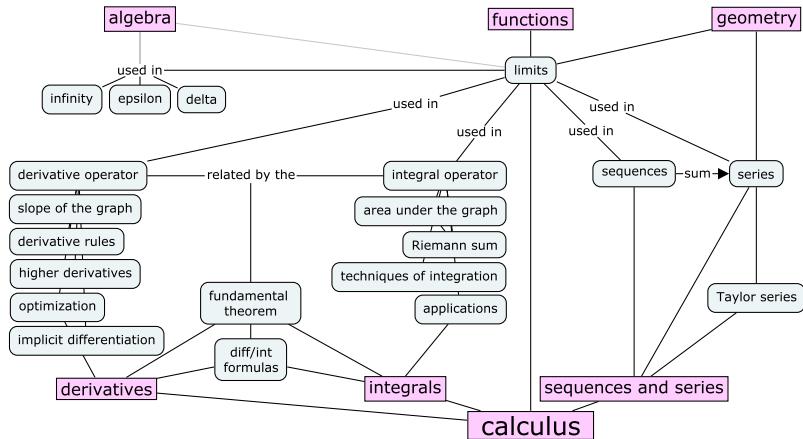

Figure 5.1: The main topics in calculus are limits, derivatives, integrals, sequences, and series. Understanding these notions and how they relate will equip you with many practical problem-solving skills.

Calculus requires a higher level of abstraction than the mathematical topics discussed in Chapter 1. We began our journey through "Math Land" with the study of *numbers*. Then we learned about *functions*, which are transformations that take real numbers as inputs and produce real numbers as outputs, $f : \mathbb{R} \to \mathbb{R}$. In calculus, the derivative and integral *operators* are procedures that take functions as inputs and produce functions as outputs. Let $\{\mathbb{R} \to \mathbb{R}\}$ denote the set of all functions that take real numbers as inputs and produce real numbers as outputs. The derivative operator takes functions as inputs and produces functions as outputs:

$$\frac{d}{dx} : \{\mathbb{R} \to \mathbb{R}\} \quad \to \quad \{\mathbb{R} \to \mathbb{R}\}.$$

More specifically, the derivative operator $\frac{d}{dx}$ acts on a function $f(x)$ to produce its derivative function: $\frac{d}{dx}[f(x)] = f'(x)$.

Differential calculus

Consider the function $f(x)$, which takes real numbers as inputs and produces real numbers as outputs, $f : \mathbb{R} \to \mathbb{R}$. The input variable for the function f is usually denoted x, but we will sometimes also

use the variables u, t, and τ to denote the inputs. The function's output is denoted $f(x)$ and is usually identified with the y-coordinate in graphs.

The *derivative* function, denoted $f'(x)$, $\frac{d}{dx}f(x)$, $\frac{df}{dx}$, or $\frac{dy}{dx}$, describes the *rate of change* of the function $f(x)$. For example, the constant function $f(x) = c$ has derivative $f'(x) = 0$ since the function $f(x)$ does not change at all.

The derivative function describes the *slope* of the graph of the function $f(x)$. The derivative of a line $f(x) = mx + b$ is $f'(x) = m$ since the slope of this line is equal to m. In general, the slope of a function is different at different values of x. For a given choice of input $x = x_0$, the value of the derivative function $f'(x_0)$ is equal to the slope of $f(x)$ as it passes through the point $(x_0, f(x_0))$.
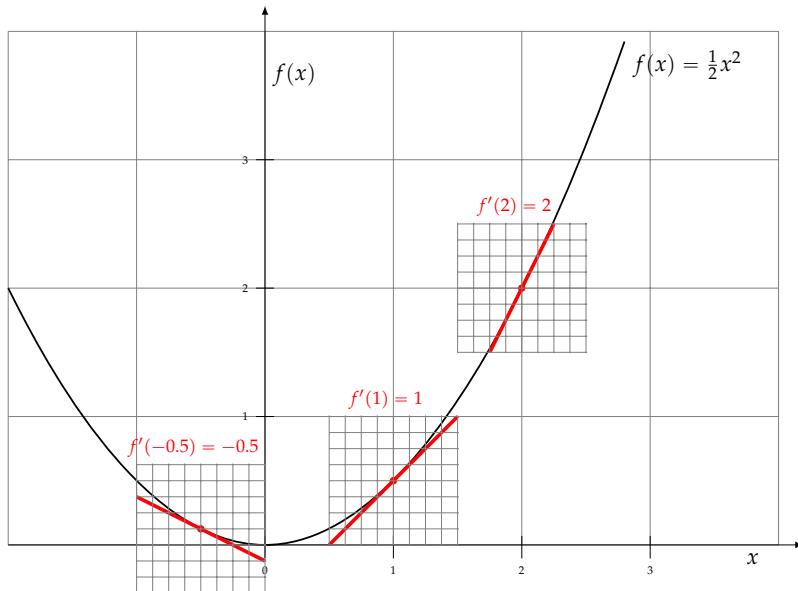

Figure 5.2: The diagram illustrates how to compute the derivative of the function $f(x) = \frac{1}{2}x^2$ at three different points on the graph of the function. To calculate the derivative of $f(x)$ at $x = 1$, we can "zoom in" near the point $(1, \frac{1}{2})$ and draw a line that has the same slope as the function. We can then calculate the slope of the line using a rise-over-run calculation, aided by the mini coordinate system that is provided. The derivative calculations for $x = -\frac{1}{2}$ and $x = 2$ are also shown. Note that the slope of the function is different for each value of x. What is the value of the derivative at $x = 0$? Is there a general pattern that describes the slope of the graph for all x?

The derivative function $f'(x)$ describes the slope of the graph of the

function $f(x)$ for all inputs $x \in \mathbb{R}$. The derivative function is a function of the form $f' : \mathbb{R} \to \mathbb{R}$. In our study of mechanics, we learned about the position function $x(t)$ and the velocity function $v(t)$, which describe the motion of an object over time. The velocity is the derivative of the object's position with respect to time $v(t) = \frac{dx}{dt} = x'(t)$.

The derivative function $f'(x)$ is a property of the original function $f(x)$. Indeed, this is where the name *derivative* comes from: $f'(x)$ is not an independent function—it is *derived* from the original function $f(x)$. In mechanics, the function $x(t)$ describes an object's position as a function of time, and the velocity function $v(t)$ describes one property of the position function, namely, how fast the object's position is changing. Similarly, the acceleration function $a(t)$ describes the rate of change of the function $v(t)$.

The *derivative operator*, denoted $\frac{d}{dx}$ or simply D, takes as input a function $f(x)$ and produces as output the derivative function $f'(x)$. The derivative operator notation is useful because it makes it clear that differentiation is an *operation* you can apply to any function to obtain its derivative:

$$ f'(x) = \frac{d}{dx} f(x). $$

The derivative operator $\frac{d}{dx}$ acts on the original function $f(x)$ to produce the derivative function $f'(x)$, which describes the rate of change of f for all x. Applying the derivative operator to a function is also called "taking the derivative" of a function.

For example, the derivative of the function $f(x) = \frac{1}{2}x^2$ is the function $f'(x) = x$. We can describe this relationship as $(\frac{1}{2}x^2)' = x$ or as $\frac{d}{dx}(\frac{1}{2}x^2) = x$. You should flip back to Figure 5.2 and use the graph to prove to yourself that the slope of $f(x) = \frac{1}{2}x^2$ is described by $f'(x) = x$ everywhere on the graph.

Differentiation techniques

We'll formally define the derivative operation in Section 5.6. Afterward, we'll develop various techniques for computing derivatives. Computing derivatives is not a complicated task once you learn how to use the derivative formulas. If you flip ahead to Section 5.7 (page 336), you'll find a table of formulas for taking the derivatives of common functions. In Section 5.8, we'll learn the general rules for computing derivatives of sums, products, and compositions of functions.

Applications of derivatives

Once you develop your ability to find derivatives, you'll be able to use this skill to perform several useful tasks.

Optimization The most prominent application of differential calculus is *optimization*: the process of finding a function's maximum and minimum values. When a function reaches its maximum value, its derivative momentarily becomes zero. The function increases just before it reaches its maximum, and the function decreases just after its maximum. At its maximum value, the function is horizontal, and $f'(x) = 0$ at this point.
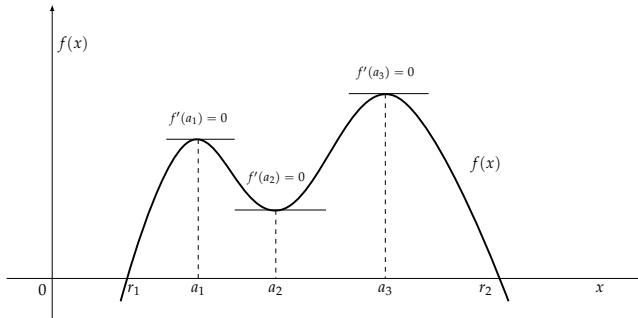

Figure 5.3: The *critical points* of a function occur where the function's derivative equals zero. The critical points of the illustrated function $f(x)$ are $x = a_1$, $x = a_2$, and $x = a_3$. You can use the critical points to find the location of a function's maxima and minima. The point $(a_1, f(a_1))$ is called a *local maximum* of the function, the point at $x = a_2$ is a *local minimum*, while the point at $x = a_3$ is the function's *global maximum*.

The values of x for which $f'(x) = 0$ are called the *critical points* of the function $f(x)$. To find the maximum of a function, we start by compiling a list of its critical points, then go through the list to find the point where the function takes on its largest value. We will discuss the details of this optimization algorithm in Section 5.10.

Tangent lines The *tangent line* to the function $f(x)$ at $x = x_0$ corresponds to the line that passes through the point $(x_0, f(x_0))$ and has the same slope as the function at that point. The word *tangent* comes from the Latin *tangere*, meaning "to touch."

The tangent line to the function $f(x)$ at the point $x = x_0$ is described by the equation

$$T_1(x) = \underbrace{f'(x_0)}_{m} x \ + \ \underbrace{(f(x_0) - f'(x_0)x_0)}_{b} = f(x_0) \ + \ f'(x_0)(x - x_0).$$

The tangent line $T_1(x)$ is an approximation to the function $f(x)$ near the coordinate $x = x_0$. The approximation $T_1(x)$ is equal to the function $f(x)$ at $x = x_0$ since the tangent line passes through the point
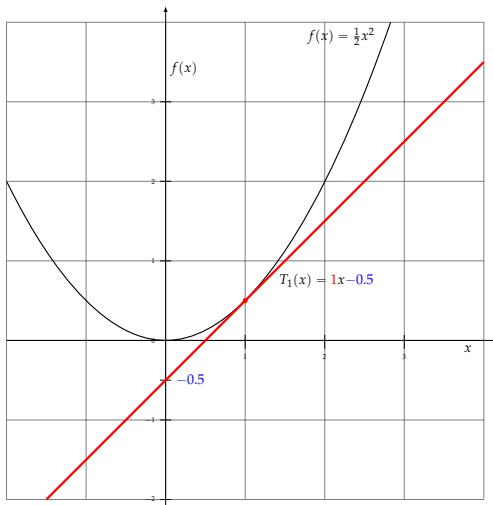
Figure 5.4: An illustration of the tangent line to the function $f(x) = \frac{1}{2}x^2$ at the point $x_0 = 1$. The equation of the tangent line is $T_1(x) = 1x - 0.5$.

$(x_0, f(x_0))$. For coordinates near $x = x_0$, the approximation is also accurate since $T_1(x)$ has the same slope as the function $f(x)$. As the input value x moves farther from x_0, the tangent becomes less accurate at approximating the function $f(x)$.

Integral calculus

The *integral* of $f(x)$ corresponds to the computation of the area under the graph of $f(x)$. The area under $f(x)$ between the points $x = a$ and $x = b$ is denoted as follows:

$$A(a,b) = \int_a^b f(x)\, dx.$$

The area $A(a,b)$ is bounded by the function $f(x)$ from above, by the x-axis from below, and by two vertical lines at $x = a$ and $x = b$. The points $x = a$ and $x = b$ are called the limits of integration. The $\int$ sign comes from the Latin word *summa*. The integral is the "sum" of the values of $f(x)$ between the two limits of integration.

The *integral function* $F(c)$ corresponds to the area calculation as a function of the upper limit of integration:

$$F(c) \overset{\text{def}}{=} \int_0^c f(x)\, dx.$$

There are two variables and one constant in this formula. The input variable c describes the upper limit of integration. The *integration*
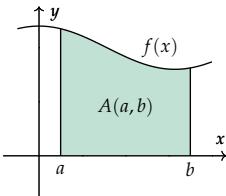
Figure 5.5: The integral of the function $f(x)$ between $x = a$ and $x = b$ corresponds to the area $A(a, b)$.

variable x performs a sweep from $x = 0$ until $x = c$. The constant 0 describes the lower limit of integration. Note that choosing $x = 0$ for the starting point of the integral function was an arbitrary choice.

The integral function $F(c)$ contains the "precomputed" information about the area under the graph of $f(x)$. Recall the derivative function $f'(x)$, which tells us the "slope of the graph" property of the function $f(x)$ for all values of x. Similarly, the integral function $F(c)$ tells us the "area under the graph" property of the function $f(x)$ for *all* possible limits of integration.

The area under $f(x)$ between $x = a$ and $x = b$ is obtained by calculating the *change* in the integral function as follows:

$$A(a, b) = \int_a^b f(x)\, dx = F(b) - F(a).$$
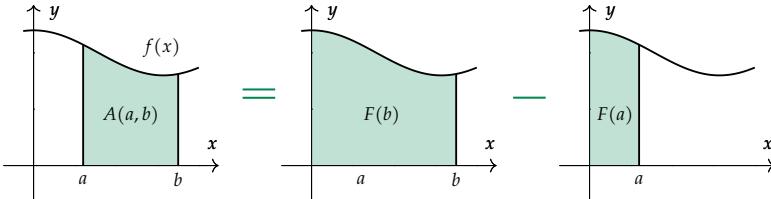
Figure 5.6: The integral function $F(x)$ computes the area under the curve $f(x)$ starting from $x = 0$. The area under $f(x)$ between $x = a$ and $x = b$ is computed using the formula $A(a, b) = F(b) - F(a)$.

Integration techniques

The bulk of the new material needed to understand integral calculus lies in learning various techniques for calculating integrals of functions. Computing integrals is not as easy as computing derivatives, because there are no general rules to follow.

In Section 5.15, we'll describe a number of common techniques for integration. These techniques will enable you to compute the integrals of polynomial functions, exponential functions, logarithmic

functions, and trigonometric functions. While these techniques will help you compute integrals in many situations, the process of computing integrals remains somewhat of an art. In art, there are no rules to follow—as an artist, you must be creative and test different approaches until you find one that works.

Applications of integration

Integral calculations have widespread applications to more areas of science than are practical to list here. Let's explore a few examples to gain a general idea of how integrals are applied in the real world.

Computing totals Integral calculations are needed every time we want to compute the total of some quantity that changes over time. If the quantity in question remains constant over time, we can multiply this quantity by the time to find the total quantity. For example, if your monthly rent is \$720, your annual rent is $R = \$720 \times 12$.

But what if your rent changes over time? Imagine a crazy landlord who demands you pay on a daily basis and changes the daily rent $r(t)$ each day. Some days rent is \$20/day, some days \$23/day, and some days he lets you stay for only \$15/day. In this situation, computing your annual rent involves the integral $R = \int_0^{365} r(t)\, dt$, which describes the calculation of the daily rate $r(t)$ times the duration of each day dt summed over all the days in the year.

Computing potentials In Section 4.6 we defined the notion of potential energy as the negative of the work done when moving an object against a conservative force. We studied two specific cases: gravitational potential energy $U_g(h) \stackrel{\text{def}}{=} -\int_0^h \vec{F}_g \cdot d\vec{y} = mgh$, and spring potential energy $U_s(x) \stackrel{\text{def}}{=} -\int_0^x \vec{F}_s(y) \cdot d\vec{y} = \frac{1}{2}kx^2$. Understanding integrals will allow you to solidify your understanding of the connection between each force $\vec{F}_?(x)$ and its associated potential energy $U_?(x)$.

Computing moments of inertia An object's moment of inertia describes how difficult it is to make the object turn. The moment of inertia is computed as the following integral:

$$ I = \int_{\text{obj}} r^2 \, dm. $$

In the mechanics chapter, I asked you to memorize the formulas for $I_{\text{disk}} = \frac{1}{2}mR^2$ and $I_{\text{sphere}} = \frac{2}{5}mR^2$ because it was not yet time to

explain the details of integral calculations. After learning about integrals, you'll be able to derive the formulas for I_{disk} and I_{sphere} on your own.

Solving differential equations   One of the most important applications of integrals is their ability to "undo" the derivative operation. Recall Newton's second law $F_{\text{net}}(t) = ma(t)$, which can also be written as

$$\frac{F_{\text{net}}(t)}{m} = a(t) = x''(t) = \frac{d}{dx}\left(\frac{d}{dx}x(t)\right).$$

In Chapter 2 we learned how to use integration to solve for $x(t)$ in special cases where the net force is constant $F_{\text{net}}(t) = F_{\text{net}}$. In this chapter, we'll revisit the procedure for finding $x(t)$, and learn how to calculate the motion of an object affected by an external force that varies over time $F_{\text{net}}(t)$.

Limits

The main new tool we'll use in calculus is the notion of a *limit*. In calculus we use limits to describe what happens to mathematical expressions when one variable becomes very large, or alternately becomes very small.

For example, to describe a situation where a number n becomes bigger and bigger, we can say,

$$\lim_{n\to\infty}(\text{expression involving } n).$$

This expression is read, "In the limit as n goes to infinity, expression involving n."

Another type of limit occurs when a small, positive number—for example $\delta > 0$, the Greek letter *delta*—becomes progressively smaller and smaller. The precise mathematical statement that describes what happens when the number δ tends to 0 is

$$\lim_{\delta\to 0}(\text{expression involving } \delta),$$

which reads, "In the limit as δ goes to zero, expression involving δ."

Derivative and integral operations are both defined in terms of limits, so understanding limits is essential for calculus. We'll explore limits in more detail and discuss their properties in Section 5.4.

Sequences

So far, we've studied functions defined for real-valued inputs $x \in \mathbb{R}$. We can also study functions defined for natural number inputs $n \in \mathbb{N}$. These functions are called *sequences*.

A sequence is a function of the form $a : \mathbb{N} \to \mathbb{R}$. The sequence's input variable is usually denoted n or k, and it corresponds to the *index* or number in the sequence. We describe sequences either by specifying the formula for the nth term in the sequence or by listing all the values of the sequence:

$$a_n, n \in \mathbb{N} \quad \Leftrightarrow \quad (a_0, a_1, a_2, a_3, a_4, \dots).$$

Note the new notation for the input variable as a subscript. This is the standard notation for describing sequences. Also note the sequence continues indefinitely.

An example of a sequence is

$$a_n = \frac{1}{n^2}, \; n \in \mathbb{N}^* \quad \Leftrightarrow \quad \left(\frac{1}{1}, \frac{1}{4}, \frac{1}{9}, \frac{1}{16}, \frac{1}{25}, \dots \right).$$

This sequence is only defined for strictly positive natural numbers $\mathbb{N}^* = \{1, 2, 3, 4, \dots\}$ as the input $n = 0$ yields a divide-by-zero error.

The fundamental question we can ask about sequences is whether they *converge* in the limit when n goes to infinity. For instance, the sequence $a_n = \frac{1}{n^2}$ converges to 0 as n goes to infinity. We can express this fact with the limit expression $\lim_{n \to \infty} \frac{1}{n^2} = 0$.

We'll discuss sequences in more detail in Section 5.18.

Series

Suppose we're given a sequence a_n and we want to compute the sum of all the values in this sequence.

To describe the sum of the 3rd, the 4th, and the 5th elements in the sequence a_n, we turn to summation notation:

$$a_3 + a_4 + a_5 = \sum_{3 \leqslant n \leqslant 5} a_n = \sum_{n=3}^{5} a_n.$$

The capital Greek letter *sigma* stands in for the word *sum*, and the range of index values included in this sum is denoted below and above the summation sign.

The partial sum of the sequence values a_n ranging from $n = 0$ until $n = N$ is denoted as

$$S_N = \sum_{n=0}^{N} a_n = a_0 + a_1 + a_2 + \cdots + a_{N-1} + a_N.$$

The *series* $\sum a_n$ is the sum of *all* the values in the sequence a_n:

$$\sum a_n = S_\infty = \lim_{N \to \infty} S_N = \sum_{n=0}^{\infty} a_n = a_0 + a_1 + a_2 + a_3 + a_4 + \cdots.$$

Note this is an infinite sum.

Series techniques

The main mathematical question we'll study with series is the question of their convergence. We say a series $\sum a_n$ *converges* if the infinite sum $S_\infty = \sum_{n\in\mathbb{N}} a_n$ equals some finite number $L \in \mathbb{R}$:

$$S_\infty = \sum_{n=0}^{\infty} a_n = L \quad \Rightarrow \quad \text{the series } \sum a_n \text{ converges.}$$

We call L the *limit* of the series $\sum a_n$. If the infinite sum $S_\infty = \sum_{n\in\mathbb{N}} a_n$ does not converge, we say the series $\sum a_n$ *diverges*. Examples of divergent series include series that "blow up" to infinity or negative infinity, or series that oscillate between different values and fail to "settle down" close to a single value L.

   The main series technique you need to learn is how to spot the differences between series that converge and series that diverge. You'll learn how to perform different *convergence tests* on the terms in the series, which will indicate whether the infinite sum converges or diverges.

Applications

Series are a powerful computational tool. We can use series to compute approximations to numbers and functions.

   For example, the number e can be computed as the following series:

$$e = \sum_{n=0}^{\infty} \frac{1}{n!} = 1 + 1 + \frac{1}{2\cdot 1} + \frac{1}{3\cdot 2} + \frac{1}{4\cdot 3\cdot 2} + \frac{1}{5\cdot 4\cdot 3\cdot 2} + \cdots.$$

The factorial operation $n!$ is the product of n times all integers smaller than n: $n! = n(n-1)(n-2)\cdots 3\cdot 2\cdot 1$. As we compute more terms from the series, our estimate of the number e becomes more accurate. The partial sum of the first six terms (as shown above) gives us an approximation of e that is accurate to three decimals. The partial sum of the first 12 terms gives us e to an accuracy of nine decimals.

   Another useful thing you can do with series is approximate functions by infinitely long polynomials. The *power series* approximation for a function $f(x)$ is defined as the series

$$f(x) = \sum_{n=0}^{\infty} c_n x^n = c_0 + c_1 x + c_2 x^2 + c_3 x^3 + c_4 x^4 + \cdots.$$

Each term in the series is of the form $a_n = c_n x^n$, where c_n is a constant that depends on the function $f(x)$.

For example, the power series of $\sin(x)$ is

$$\sin(x) = \overbrace{x}^{T_1(x)} \underbrace{- \frac{x^3}{3!} + \frac{x^5}{5!}}_{T_5(x)} - \frac{x^7}{7!} + \frac{x^9}{9!} - \frac{x^{11}}{11!} + \cdots.$$

We can truncate the infinite series anywhere to obtain an approximation to the function. The function $T_5(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!}$ is the best approximation to the function $\sin(x)$ by a polynomial of degree five. The equation of the tangent line $T_1(x)$ at $x = 0$ is a special case of the power series approximation procedure, which approximates the function as a first-degree polynomial.

We'll continue the discussion on series, their properties, and their applications in Section 5.19.

If you haven't noticed yet from glancing at the examples so far, the common theme underpinning all the topics of calculus is the notion of *infinity*. We now turn our attention to the infinite.

5.3   Infinity

Working with infinitely small quantities and infinitely large quantities can be tricky business. It's important that we develop the appropriate language for describing these concepts as soon as possible. Like, now.

Infinitely large

The number ∞ is *really* large. How large? Larger than any number you can think of. Think of any number n. It is true that $n < \infty$. Now think of a bigger number N. It will still hold true that $N < \infty$. In fact, any finite number you can think of, no matter how large, will always be less than ∞.

Technically speaking, ∞ is not a number; infinity is a *process*. You can think of ∞ as the answer you obtain by starting from 0 and continuously adding 1 *forever*.

To see why $N < \infty$ for any finite number N, consider the following reasoning. When we add 1 to a number, we obtain a larger number. The operation $+1$ is equivalent to taking one unit step to the right on the number line. For any n, it is true that $n < n + 1$. To get to infinity we start from $n = 0$ and keep adding 1. After N steps, we'll arrive at $n = N$. But then we must continue adding 1 and obtain $N + 1$, $N + 2$, $N + 3$, and so on. Since adding 1 always creates a

Even if the pieces of string are *infinitely small*, because there are *infinitely many* of them, they still add to ℓ.

The take-home message is that as long as you clearly define your limits, you can use infinitely small numbers in your calculations. The notion of a limit is one of the central ideas in this course.

Limits at infinity

In math, we're often interested in describing what happens to a certain function when its input variable tends to infinity. This information helps us draw the function's graph. Does $f(x)$ approach a finite number, or does it keep on growing to ∞?

As an example of this type of calculation, consider the limit of the function $f(x) = \frac{1}{x}$ as x goes to infinity:

$$\lim_{x \to \infty} f(x) = \lim_{x \to \infty} \frac{1}{x} = 0.$$

This statement is true, even though the function $\frac{1}{x}$ never *actually* reaches zero. The function gets closer and closer to the x-axis but never touches it. This is why the concept of a limit is useful: it allows us to write $\lim_{x \to \infty} f(x) = 0$ even though $f(x) \neq 0$ for any $x \in \mathbb{R}$.

The function $f(x)$ is said to *converge* to the number L if the function approaches the value L for large values of x:

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

We say "The limit of $f(x)$ as x goes to infinity is the number L." See Figure 5.7 for an illustration. The limit expression is a concise way of saying the following precise mathematical statement: for *any* precision $\epsilon > 0$, there exists a starting point S, after which $f(x)$ equals L within a precision ϵ.
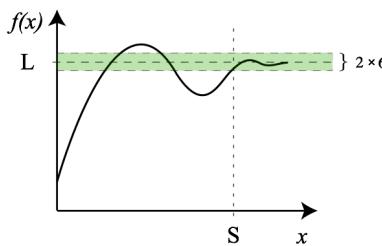

Figure 5.7: A function $f(x)$ whose oscillations around the value L are smaller than ϵ for all $x \geqslant S$. The starting point S depends on the choice of precision ϵ.

5.6   Derivatives

In the beginning of the chapter we introduced the derivative concept by identifying the derivative with the slope of the function's graph. This graphical representation of derivatives and the intuition that comes with it are very important: this is how mathematicians and physicists usually "think" about derivatives. It is equally important to understand the formal definition of the derivative operation, so this is what we will cover next. Afterward, we'll build some practical skills for calculating derivatives of functions.

Definition

The *derivative* of a function is defined as

$$f'(x) \overset{\text{def}}{=} \lim_{\delta \to 0} \frac{f(x + \delta) - f(x)}{\delta}.$$

The definition of the derivative comes from the rise-over-run formula for calculating the slope of a line:

$$\frac{\text{rise}}{\text{run}} = \frac{\Delta y}{\Delta x} = \frac{y_f - y_i}{x_f - x_i} = \frac{f(x + \delta) - f(x)}{x + \delta - x}.$$

By making δ tend to zero in the above expression, we are able to obtain the slope of the function $f(x)$ at the point x.

Derivatives occur so often in math that people have devised many ways to denote them. Don't be fooled by this multitude of notations—all of them refer to the same concept:

$$Df(x) = f'(x) = \frac{d}{dx}f(x) = \frac{df}{dx} = \nabla f.$$

Example   Let's calculate the derivative of $f(x) = 2x^2 + 3$ to illustrate how the complicated-looking derivative formula works:

$$f'(x) = \lim_{\delta \to 0} \frac{f(x + \delta) - f(x)}{\delta} = \lim_{\delta \to 0} \frac{2(x + \delta)^2 + 3 - (2x^2 + 3)}{\delta}.$$

We can simplify the fraction inside the limit:

$$\frac{2x^2 + 4x\delta + 2\delta^2 - 2x^2}{\delta} = \frac{4x\delta + 2\delta^2}{\delta} = \frac{4x\delta}{\delta} + \frac{2\delta^2}{\delta} = 4x + 2\delta.$$

The second term of this expression disappears when we take the limit to obtain the final answer:

$$f'(x) = \lim_{\delta \to 0} (4x + 2\delta) = 4x + 0 = 4x.$$

Product rule

The derivative of a product of two functions is obtained as follows:

$$[f(x)g(x)]' = f'(x)g(x) + f(x)g'(x).$$

Quotient rule

The *quotient rule* tells us how to obtain the derivative of a fraction of two functions:

$$\left[\frac{f(x)}{g(x)}\right]' = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}.$$

Chain rule

If you encounter a situation that includes an inner function and an outer function, like $f(g(x))$, you can obtain the derivative by a two-step process:

$$[f(g(x))]' = f'(g(x))g'(x).$$

In the first step, leave the inner function $g(x)$ alone, and focus on taking the derivative of the outer function $f(x)$. This step gives us $f'(g(x))$, the value of f' evaluated at $g(x)$. As the second step, we multiply the resulting expression by the derivative of the *inner* function $g'(x)$.

The chain rule tells us the derivative of a composite function is calculated as the product of the derivative of the outer function and the derivative of the inner function.

Simple example   Consider the following derivative calculation:

$$\left[\sin(x^2)\right]' = \cos(x^2)\left[x^2\right]' = \cos(x^2)2x.$$

The chain rule also applies to functions of functions of functions $f(g(h(x)))$. To take the derivative, start from the outermost function and work your way toward x.

$$[f(g(h(x)))]' = f'(g(h(x)))g'(h(x))h'(x).$$

stealing of natural resources from Third World countries, optimize anything that moves, basically. Therefore, the System wants *you*—the young and powerful generation of the future—to learn this important skill and become faithful employees of corporations. The corporates want you to learn calculus so you can help them optimize things, ensuring the smooth continuation of the whole enterprise.

Mathematical knowledge does not come with an ethics manual to help you decide what should and should not be optimized; this responsibility falls on you. If, like me, you don't want to become a corporate sellout, you can always choose to use calculus for science. It doesn't matter whether it will be physics, medicine, or running your own company, it is all good. Just stay away from the System. Please do this for yourself and our future, will you?

Having said these words of warning, let's now proceed so I can show you the powerful optimization algorithm.

5.10 Optimization algorithm

This section shows and explains the details of the algorithm for finding the maximum of a function. This is called *optimization*, as in finding the optimal solution to a problem.

Say you have the function $f(x)$, which represents a real-world phenomenon. For example, $f(x)$ could represent how much *fun* you have as a function of alcohol consumed during one evening. We all know that with too much x, the fun stops and you find yourself, as the Irish say, "talking to God on the big white phone." Too little x and you might not have enough Dutch courage to chat up that girl/guy from the table across the room. To have as much fun as possible, you want to find the alcohol consumption x^* where f takes on its maximum value.

This is one of the prominent applications of calculus (I'm talking about optimization, not alcohol consumption). This is why you've been learning about all those limits, derivative formulas, and differentiation rules in previous sections.

Definitions

- x: the *variable* we can control
- $[x_i, x_f]$: the interval of values from which x can be chosen. The values of x must obey $x_i \leqslant x \leqslant x_f$. These are the *constraints* of the optimization problem. For the drinking optimization problem, $x \geqslant 0$ since you can't drink negative alcohol, and probably

Discussion

We have reached the half-point of the calculus chapter. We learned about derivatives and described applications of derivatives to optimization problems, finding tangent lines, related rates, etc.

Before you continue reading about integrals in the second half of the chapter, I highly recommend you attempt to solve some of the derivative problems starting on page 431.

Another thing I would recommend is to watch some of the lectures of the *Highlights of Calculus* course by Prof. Gilbert Strang (see `https://youtube.com/playlist?list=PLBE9407EA64E2C318`).

5.12   Integrals

We now begin our discussion of integrals, the second topic in calculus. An integral is a fancy way of computing the area under the graph of a function. Integral calculus is usually taught as a separate course after differential calculus, but this separation can be counterproductive. The easiest way to understand integration is to think of it as the inverse of the derivative operation. Integrals are antiderivatives. Once you realize this fundamental fact, you'll be able to apply all your differential calculus knowledge to the domain of integral calculus. In differential calculus, we learned how to take a function $f(x)$ and find its derivative $f'(x)$. In integral calculus, we'll be given a function $f(x)$ and we'll be asked to find its *antiderivative* function $F(x)$. The antiderivative of $f(x)$ is a function $F(x)$ whose derivative equals $f(x)$.

In this section, we'll learn about two tasks: how to compute antiderivatives, and how to compute the area under the graph of $f(x)$. Confusingly, both of these operations are called *integration*. To avoid any possibility of confusion, we'll define the two concepts right away:

- The *indefinite integral* of $f(x)$ is denoted $\int f(x)dx = F(x) + C$. To compute the indefinite integral of $f(x)$, you must find a function $F : \mathbb{R} \to \mathbb{R}$, such that $F'(x) = f(x)$. The indefinite integral is the antiderivative function.

- The *definite integral* of $f(x)$ between $x = a$ and $x = b$ is denoted $\int_a^b f(x)dx = A(a,b)$. Definite integrals correspond to the computation of the area under the function $f(x)$ between $x = a$ and $x = b$. The definite integral is a number $A(a,b) \in \mathbb{R}$.

The two integration operations are related. The area under the curve $A(a,b)$ can be computed as the *change* in the antiderivative function, using to the formula $A(a,b) = \big[F(x) + C\big]_a^b = F(b) - F(a)$.

Definitions

You should already be familiar with these concepts:

- $\mathbb{R}$: the set of real numbers
- $f(x)$: a function of the form $f : \mathbb{R} \to \mathbb{R}$, which means f takes real numbers as inputs and produces real numbers as outputs
- $\lim_{\delta \to 0}$: a limit expression in which the number δ tends to zero
- $f'(x)$: the derivative of $f(x)$ is the rate of change of f at x:

$$f'(x) \overset{\text{def}}{=} \lim_{\delta \to 0} \frac{f(x + \delta) - f(x)}{\delta}.$$

  The derivative is a function of the form $f' : \mathbb{R} \to \mathbb{R}$.

These are the new concepts, which we will learn about in integral calculus:

- $A(a, b)$: the value of the *area* under the curve $f(x)$ from $x = a$ until $x = b$. The area $A(a, b)$ is computed as the following integral

$$A(a, b) = \int_a^b f(x)\, dx.$$

  The $\int$ sign stands for *sum*. Indeed, the integral is the "sum" of $f(x)$ for all values of x between a and b.

- $A_0(x)$: the *integral function* of $f(x)$. The integral function corresponds to the computation of the area under $f(x)$ as a function of the upper limit of integration:

$$A_0(x) \overset{\text{def}}{=} A(0, x) = \int_0^x f(u)\, du.$$

  The choice of $x = 0$ as the lower limit of integration is arbitrary.

- $F(x) + C$: The *antiderivative* function of the function $f(x)$. An antiderivative function is defined as a function whose derivative equals to $f(x)$. The antiderivative function always includes an additive constant C. If the function $F(x)$ is an antiderivative (obeys $F'(x) = f(x)$) then the function $F(x) + C$ is also an antiderivative since

$$\frac{d}{dx}[F(x) + C] = f(x),$$

  for any constant C.

- The fundamental theorem of calculus (FTC) states that the integral function $A_0(x)$ is equal to the antiderivative function $F(x)$ up to an additive constant C:

$$A(0, x) = A_0(x) \stackrel{\text{FTC}}{=} F(x) + C.$$

The fundamental theorem leads us to the following formula for computing the area $A(a, b)$:

$$A(a, b) = A(0, b) - A(0, a) = A_0(b) - A_0(a) = F(b) - F(a).$$

The area under the curve, $A(a, b)$, is equal to the change in the antiderivative function $F(x)$ between $x = a$ and $x = b$.

The area under the curve

An integral describes the computation of the area under the curve $f(x)$ between $x = a$ and $x = b$:

$$A(a, b) = \int_a^b f(x)\, dx.$$

We refer to the numbers a and b as the *limits of integration*. The location where the integral starts, $x = a$, is called the *lower limit* of integration. The location where the integral stops, $x = b$, is called the *upper limit* of integration.
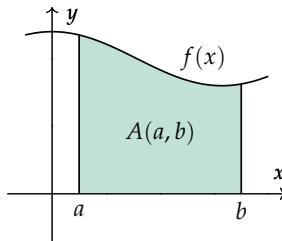

Figure 5.16: The integral of the function $f(x)$ between $x = a$ and $x = b$ corresponds to the area $A(a, b)$.

The integral as a function

The *integral function* of $f(x)$ describes the "running total" of the area under the curve $f(x)$ as a function of the upper limit of integration:

$$A_0(x) = A(0, x) = \int_0^x f(u)\, du.$$

5.13 Riemann sums

Our discussion in the previous section focused on the inverse relationship between the integral operator $\int f(x)dx$ and the derivative operator $\frac{d}{dx}$. We learned the antiderivative function $F(x)$ can be used to compute the area under a curve $f(x)$ using the formula $\int_a^b f(x)\,dx = F(b) - F(a)$. Thus, with your differentiation skills and some reverse engineering, you can now handle integrals too.

Is there a way to compute integrals without referring to the derivative operation? No course on calculus would be complete without telling the classic "rectangles story" for computing definite integrals, which goes by the name *Riemann sum*. The Riemann sum is a procedure for computing the area under a curve by breaking up the area into many, little, rectangular strips with heights that vary according to $f(x)$. To obtain the total area under the curve, we sum all the areas of these little rectangles.

First, like a cast of characters, we'll introduce some definitions.

Definitions

- $f(x)$: a function $f : \mathbb{R} \to \mathbb{R}$
- a: where the sum starts
- b: where the sum stops
- $A(a, b)$: the exact value of the area under the curve $f(x)$ from $x = a$ until $x = b$
- $S_n(a, b)$: an approximation to the area A in terms of n rectangles
- s_k: the area of the kth rectangle when counting from the left

In this section we will consider the calculation of the area under the curve $f(x) = x^3 - 5x^2 + x + 10$ between $x = -1$ and $x = 4$. Figure 5.18 shows the graph of $f(x)$ and an approximation of the area under the curve as the sum of the areas of 12 rectangles.

Formulas

The combined-area approximation is given by the *sum* of the areas of the little rectangles:

$$S_n(a, b) = \sum_{k=1}^{n} s_k.$$

Each of the little rectangles has an area s_k given by the rectangle's *height* multiplied by its *width*. The height of each rectangle will vary, but the width is constant. Why constant? Riemann figured that having each rectangle with a constant width Δx would make it easy to
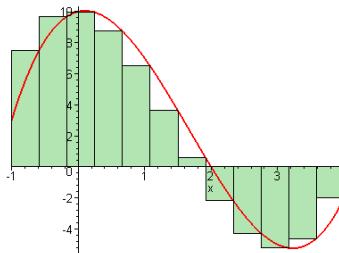
Figure 5.18: An approximation of the area under the function $f(x) = x^3 - 5x^2 + x + 10$ between $x = -1$ and $x = 4$ using $n = 12$ rectangles.

calculate the approximation. The total length of the interval from $x = a$ to $x = b$ is $(b - a)$. When we divide this length into n equally spaced segments, each segment will have width Δx given by

$$\Delta x = \frac{b - a}{n}.$$

Okay, we have the width formula; now let's find the height of the kth rectangle in the sequence of rectangles. For the rectangles, we pick isolated "samples" of $f(x)$ for the following values:

$$x_k = a + k\Delta x, \text{ for } k \in \{1, 2, 3, \ldots, n\},$$

with all rectangles equally spaced $\Delta x = \frac{b-a}{n}$ apart.

The function's height varies as we move along the x-axis. The area of each rectangle is equal to its height $f(x_k)$ times its width:

$$s_k = f(a + k\Delta x)\Delta x.$$

Now, my dear students, I want you to stare at the above equation and do some simple calculations to check that you understand. There is no point in continuing if you are just taking my word for it. **Verify** that when $k = 1$, the formula gives the area of the first little rectangle. **Verify also** that when $k = n$, the formula $x_k = a + k\Delta x$ reaches the upper limit b.

Let's put our formula for s_k in the sum where it belongs. The Riemann sum approximation using n rectangles is given by

$$S_n(a, b) = \sum_{k=1}^{n} f(a + k\Delta x)\Delta x,$$

where $\Delta x = \frac{b-a}{n}$.

The integral is defined as the limit of the Riemann sum as n goes to infinity: $\int_a^b f(x)\, dx \overset{\text{def}}{=} \lim_{n \to \infty} S_n(a, b) = \lim_{n \to \infty} \sum_{k=1}^{n} f(a + k\Delta x)\Delta x.$

A priori, there is no reason to suspect the integral function would be related to the derivative operation. The integral corresponds to the computation of an area, whereas the derivative operation computes the slope of a function. The fundamental theorem of calculus describes the relationship between derivatives and integrals.

Theorem (fundamental theorem of calculus). *Let $f(x)$ be a continuous function on the interval $[a, b]$, and let $c \in \mathbb{R}$ be a constant. Define the function $A_c(x)$ as follows:*

$$A_c(x) \stackrel{def}{=} \int_c^x f(u) \, du.$$

Then, the derivative of $A_c(x)$ is equal to $f(x)$:

$$\frac{d}{dx}[A_c(x)] = f(x),$$

for any $x \in (a, b)$.

The fundamental theorem of calculus establishes an equivalence between the set of integral functions and the set of antiderivative functions:

$$A_c(x) = F(x) + C.$$

All integral functions $A_c(x)$ are antiderivatives of $f(x)$.

Differential calculus and integral calculus are two sides of the same coin. If you understand why the theorem is true, you will understand something very deep about calculus. Differentiation is the inverse operation of integration. Given a function $G(x) = \int g(x)dx$, we can obtain the function $g(x)$ by taking the derivative of $G(x)$: $G'(x) = g(x)$. The inverse relationship works in the other direction as well. If you're given the derivative $h'(x)$ of some unknown function $h(x)$, you can find the function $h(x)$ (up to a constant), using integration: $h(x) + C = \int h'(x)dx$.

Got proof?

There is an unspoken rule in mathematics: when the word *theorem* appears in writing, it must be followed by the word *proof*. We therefore need to look into the proof of the fundamental theorem of calculus (FTC). It is not so important you understand the details of the proof, but I still recommend you read this subsection for your general math knowledge. If you are in a rush though, feel free to skip ahead.

5.15   Techniques of integration

The operation of "taking the integral" of some function is usually much more complicated than that of taking the derivative. You can take the derivative of *any* function—no matter how complex—by using the product rule, the chain rule, and the derivative formulas. This is not true for integrals.

Plenty of integrals have no *closed-form solution*, meaning the function has no antiderivative. There is no simple procedure to follow such that you input a function and "turn the crank" until the integral comes out. Integration is a bit of an art.

Which functions *can* we integrate, and how? Back in the day, scientists collected big tables with integral formulas for various complicated functions. We can use these tables to *look up* a specific integral formula. Such table is given on page 507 in the back of the book.

We can also learn some *integration techniques* to help make complicated integrals simpler. Think of the techniques presented in this section as *adapters*. You can reach for these adapters when the function you need to integrate doesn't appear in your table of integrals, but a similar one is found in the table.

A note to all our students in the audience who are taking an integral calculus course. These integration techniques are exactly the skills you'll be expected to demonstrate on the final. Instead of using the table of integrals to look up complicated integrals, you'll need to know how to fill in the table.

For people interested in learning physics, I'll honestly tell you that if you skip this next section you won't miss much. You should read the important section on *substitution*, but there's no need to read the details of all the recipes for integrating things. For most intents and purposes, once you understand what an integral is, you can use a computer to calculate it. A good tool for calculating integrals is the computer algebra system at `live.sympy.org`.

```
>>> integrate( sin(x) )
-cos(x)
>>> integrate( x**2*exp(x) )
x**2*exp(x) - 2*x*exp(x) + 2*exp(x)
```

You can use `SymPy` for all your integration needs.

A comment to those of you reading this book for fun, without the added stress of homework and exams. Consider the next dozen pages as an ethnographic snapshot of the daily life of the undergraduate experience in science. Try to visualize the life of first-year science students, busy integrating things they don't want to integrate for many, long hours. Picture some unlucky science student locked
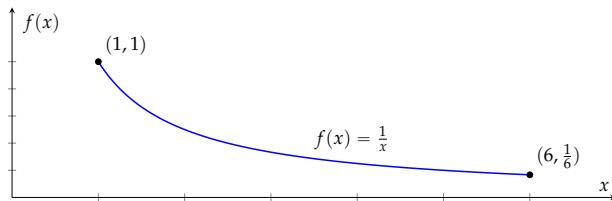
Figure 5.28: Graph of $f(x) = \frac{1}{x}$ between $x = 1$ and $x = 6$. What surface will be produced if we rotate this curve around the x-axis?
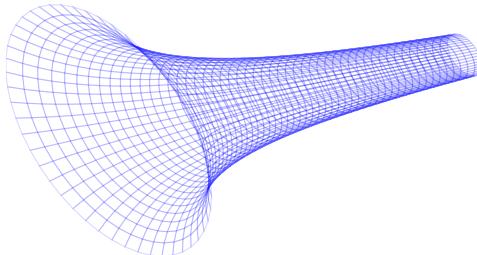

Figure 5.29: The surface of revolution generated by the section of the curve $f(x) = \frac{1}{x}$ between $x = 1$ and $x = 6$ when rotated around the x-axis.

The area of the surface of revolution traced by the graph of $f(x)$ between $x = a$ and $x = b$ as it rotates around the x-axis is given by the formula

$$A = \int_a^b 2\pi f(x)d\ell = \int_a^b 2\pi f(x)\sqrt{1 + (f'(x))^2}\, dx.$$

Volumes of revolution

Let's move on to three-dimensional integrals, or integrals over volumes. Again, we'll use the circular symmetry of the object's volume to split the object into little "pieces of volume" and then compute an integral to find the total volume.

Disk method

We can describe the volume of an object with circular symmetry as the sum of a number of disks. Each disk will have thickness dx and a radius proportional to the function $f(x)$. In other words, the function $f(x)$ describes the object's outer boundary. The area of each disk is $\pi(f(x))^2$ and its thickness is dx.

The volume of a solid of revolution with boundary $f(x)$ rotated

Ratio convergence

The numbers in the Fibonacci sequence grow indefinitely large
($\lim\limits_{n\to\infty} a_n = \infty$), while the ratio of $\frac{a_{n+1}}{a_n}$ converges to a constant:

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = \varphi = \frac{1 + \sqrt{5}}{2} \approx 1.618033\ldots$$

This constant is known as the *golden ratio*.

Calculus on sequences

If a sequence a_n is like a function $f(x)$, we should be able to perform
calculus on it. We already saw how we can take limits of sequences,
but can we also compute derivatives and integrals of sequences?

The usual derivative notion is a no-go since it depends on the
function $f(x)$ being *continuous*, and sequences are only defined for
integer values. Instead of derivatives, we can compute *finite differ-
ences*, which are sequences obtained by subtracting adjacent terms in
the sequence. Given the sequence $(a_0, a_1, a_2, a_3, a_4, \ldots)$, the first dif-
ferences sequence is the sequence $(a_1 - a_0, a_2 - a_1, a_3 - a_2, \ldots)$. Finite
differences play a big role in the study of differential equations.

We can also compute integrals of sequences, and this is the sub-
ject of the next section.

5.19   Series

Can you compute $\ln(2)$ using only a basic calculator with four oper-
ations, $\boxed{+}$, $\boxed{-}$, $\boxed{\times}$, and $\boxed{\div}$? I can tell you one way to do this;
compute the following infinite sum:

$$\ln(2) = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \cdots.$$

Since the sum is infinite, it will take a while to obtain the value of
$\ln(2)$, but if you keep adding more terms in the sum, you can com-
pute the answer $\ln(2) = 0.69314718\ldots$ to *any* precision.

Let's make the computer carry out the summation for us. First we
define the formula for the nth term in the series $a_n = \frac{(-1)^{n+1}}{n}$, then
we compute the sum of the first 100, 1000, and 1000000 terms:

```
>>> def an_ln2(n): return (-1.0)**(n+1)/n
>>> sum([ an_ln2(n)  for n in range(1,100) ])
0.698172179310195
```

Exact sums

Formulas exist for calculating the exact sum of certain series.

The sum of the geometric series of length N is

$$S_N = \sum_{n=0}^{N} r^n = 1 + r + r^2 + \cdots + r^N = \frac{1 - r^{N+1}}{1 - r}.$$

If $|r| < 1$, taking the limit $N \to \infty$ in the above expression leads to

$$S_\infty = \lim_{N \to \infty} S_N = \sum_{n=0}^{\infty} r^n = 1 + r + r^2 + r^3 + \cdots = \frac{1}{1 - r}.$$

Example　Consider the geometric series with $r = \frac{1}{2}$. Applying the above formula, we obtain

$$S_\infty = \sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \frac{1}{32} + \cdots = \frac{1}{1 - \frac{1}{2}} = 2.$$

You can visualize this infinite summation graphically. Imagine starting with a piece of paper of size one-by-one, then adding next to it a second piece of paper with half the size of the first, and a third piece with half the size of the second, and so on. The total area occupied by these pieces of papers is shown in Figure 5.32.
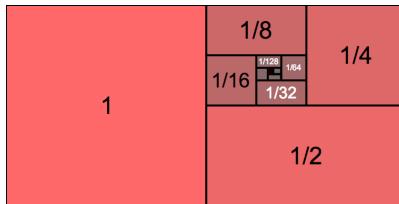

Figure 5.32: A graphical representation of the infinite sum of the geometric series with $r = \frac{1}{2}$. The area of each region corresponds to one of the terms in the series. The total area is equal to $\sum_{n=0}^{\infty} \left(\frac{1}{2}\right)^n = \frac{1}{1 - \frac{1}{2}} = 2$.

We'll now state without proof a number of other formulas where the sum of a series can be obtained as a closed-form expression.

The formulas for the sum of the first N positive integers and the sum of the squares of the first N positive integers are

$$\sum_{n=1}^{N} n = \frac{N(N + 1)}{2}, \qquad \sum_{n=1}^{N} n^2 = \frac{N(N + 1)(2N + 1)}{6}.$$

5.21 Calculus problems

In this chapter we learned about derivatives and integrals, which are mathematical operations relating to the slope of a function and the area under the graph of a function. We also learned about limits, sequences, and series. It's now time to see how much you've really learned by trying to solve some calculus problems.

Calculus hasn't changed much in the last hundred years. It is testament to this fact that many of the problems presented here were adapted from the book "Calculus Made Easy" by Silvanus Thompson, originally published[1] in 1910. These problems remain as pertinent and interesting today as they were 100 years ago.

As much as calculus is about understanding things conceptually and seeing the big picture (abstraction), calculus is also about practice. There are more than 120 problems to solve in this section. The goal is to turn differentiation and integration into routine operations that you can carry out without stressing out. You should vanquish as many problems as you need to feel comfortable with the procedures of calculus.

Okay, enough prep talk. Let's get to the problems!

Limits problems

P5.1 Use the graph of the function $f(x)$ shown in Figure 5.33 to calculate the following limit expressions:

(1) $\lim\limits_{x\to-5^-} f(x)$ (2) $\lim\limits_{x\to-5^+} f(x)$ (3) $\lim\limits_{x\to-5} f(x)$

(4) $\lim\limits_{x\to2^-} f(x)$ (5) $\lim\limits_{x\to2^+} f(x)$ (6) $\lim\limits_{x\to2} f(x)$

(7) $\lim\limits_{x\to5^-} f(x)$ (8) $\lim\limits_{x\to5^+} f(x)$ (9) $\lim\limits_{x\to5} f(x)$

(10) Is the function $f(x)$ continuous at $x = 5$?

(11) What are the intervals where the function $f(x)$ is continuous?

P5.2 Find the value of the following limit expressions:

(a) $\lim\limits_{x\to3} 4$ (b) $\lim\limits_{x\to3} 2x$ (c) $\lim\limits_{x\to3} x^2 - 2x + 2$

P5.3 Prove the following limit statement by constructing an ϵ, δ-proof:

$$\lim_{x\to5} 3x = 15.$$

Recall the ϵ, δ-game: one player (the skeptic) specifies the required precision $\epsilon > 0$, and the other player (the prover) must find a value $\delta > 0$ such that $|3x - 15| \leqslant \epsilon$ for all x in the interval $(5 - \delta, 5 + \delta)$.
Hint: Choose δ to be a multiple of ϵ.

---

[1] Full text is available at `http:/gutenberg.org/ebooks/33283` (public domain).

End matter

Conclusion

We managed to cover a lot of ground, explaining many topics and concepts in a relatively small textbook. We reviewed high school math and learned about mechanics and calculus. Above all, we examined math and physics material in an integrated manner.

If you liked or hated this book, be sure to send me feedback. Feedback is crucial so I know how to adjust the writing, the content, and the attitude of the book for future learners of math. Please take the time to drop me a line if you find a mistake or to let me know what you thought. You can reach me by email at `ivan@minireference.com`.

If you want to learn about other books in the **No Bullshit Guide** series or hear about the technology we're using at Minireference Co. to take over the textbook industry, check out the company blog at `minireference.com/blog/`. You can also find us on the twitter `@minireference` and on the facebook `fb.me/noBSguide`.

Acknowledgments

This book would not have been possible without the support and encouragement of the people around me. I am fortunate to have grown up surrounded by good people who knew the value of math and encouraged me in my studies and with this project. In this section, I want to *big up* all the people who deserve it.

First and foremost in this list are my parents from whom I have learned everything, and who have supported me throughout my life.

Next in line are all my teachers. I thank my CEGEP teachers: Karnig Bedrossian from whom I learned calculus, Paul Kenton from whom I learned how to think about physics in a chill manner, and Benoit Larose who taught me that more dimensions does not mean things get more complicated. I thank Kohur Gowrisankaran, Frank

Further reading

You have reached the end of this book, but you're only at the beginning of the journey of scientific discovery. There are a lot of cool things left for you to learn about. Below are some recommendation of subjects you might find interesting.

Electricity and Magnetism

Electrostatics is the study of the electric force $\vec{F}_e$ and the associated electric potential U_e. Here, you will also learn about the electric field $\vec{E}$ and electric potential V.

Magnetism is the study of the magnetic force $\vec{F}_b$ and the magnetic field $\vec{B}$, which are caused by electric currents flowing through wires. The current I is the total number of electrons passing through a cross-section of the wire in one second. By virtue of its motion through

space, each electron contributes to the strength of the magnetic field surrounding the wire.

The beauty of electromagnetism is that the entire theory can be described in just four equations:

$$\nabla \cdot \vec{E} = \frac{\rho}{\varepsilon_0} \qquad \text{Gauss's law}$$

$$\nabla \cdot \vec{B} = 0 \qquad \text{Gauss's law for magnetism}$$

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \qquad \text{Faraday's law of induction}$$

$$\nabla \times \vec{B} = \mu_0 \vec{J} + \mu_0 \varepsilon_0 \frac{\partial \vec{E}}{\partial t} \qquad \text{Ampère's circuital law}$$

Together, these are known as Maxwell's equations.

Vector calculus

You may be wondering what the triangle thing is. The symbol ∇ (called *nabla*) is the vector derivative operation: $\nabla \overset{\text{def}}{=} \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right)$. Guess what—you can also do calculus with vectors.

If you take a vector calculus course, you'll learn about path integrals, surface integrals, and volume integrals of vector functions. You will also learn about vector-derivatives, as well as two vector equivalents of the fundamental theorem of calculus:

- Stokes' Theorem:

$$\iint_{\Sigma} \nabla \times \vec{F} \cdot d\vec{S} = \int_{\partial \Sigma} \vec{F} \cdot d\vec{r},$$

  which states that the integral of the curl, $\text{curl}(\vec{F}) \overset{\text{def}}{=} \nabla \times \vec{F}$, of the field $\vec{F}$ over the surface Σ is equal to the circulation of $\vec{F}$ along the boundary of the surface $\partial \Sigma$.
- Gauss' Divergence Theorem:

$$\iiint_{V} \nabla \cdot \vec{F} \, dV = \iint_{\partial V} \vec{F} \cdot d\vec{S},$$

  which states the integral of the divergence, $\text{div}(\vec{F}) \overset{\text{def}}{=} \nabla \cdot \vec{F}$, of the field $\vec{F}$ over the volume V is equal to the flux of $\vec{F}$ through the volume's boundary ∂V.

Both theorems relate the total of the derivative of a quantity over some region R to the value of that quantity on the boundary of the

Appendix A

Answers and solutions

Chapter 1 solutions

Answers to exercises

E1.1 a) $x = 3$; **b)** $x = 30$; **c)** $x = 2$; **d)** $x = -3$. **E1.2 a)** $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$; **b)** $\mathbb{C}$; **c)** $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$; **d)** $\mathbb{Q}, \mathbb{R}, \mathbb{C}$; **e)** $\mathbb{R}, \mathbb{C}$. **E1.3 a)** 21; **b)** 0; **c)** $\frac{2}{27}$. **E1.4 a)** $\frac{5}{6}$; **b)** $\frac{13}{12} = 1\frac{1}{12}$; **c)** $\frac{31}{6} = 5\frac{1}{6}$. **E1.5 a)** $x = 2$; **b)** $x = 25$; **c)** $x = 100$. **E1.6 a)** $f^{-1}(x) = x^2$, $x = 16$. **b)** $g^{-1}(x) = -\frac{1}{2}\ln(x)$, $x = 0$. **E1.7 a)** $(x-1)(x-7)$; **b)** $(x+2)^2$; **c)** $(x+3)(x-3)$. **E1.8 a)** $x^2 + 2x - 15 = (x+1)^2 - 16 = 0$, which has solutions $x = 3$ and $x = -5$; **b)** $x^2 + 4x + 1 = (x+2)^2 - 3 = 0$, with solutions $x = -2 + \sqrt{3}$ and $x = -2 - \sqrt{3}$. **E1.9** $x_1 = \frac{3}{2}$ and $x_2 = -1$. **E1.10** $x = \pm\sqrt{2}$. **E1.11 a)** 8; **b)** $a^{-1}b^{-2}c^{-3} = \frac{1}{ab^2c^3}$; **c)** $8a^2$; **d)** a^6b^{-2}. **E1.12 a)** 3; **b)** 12; **c)** $\sqrt{3}$; **d)** $|a|$. **E1.13 a)** 2π; **b)** $4 + \frac{1}{4} = 4.25$; **c)** 1; **d)** x^2. **E1.14 a)** $x = \sqrt{a}$ and $x = -\sqrt{a}$; **b)** $x = \sqrt[3]{b}$; **c)** $x = \sqrt[4]{c}$ and $x = -\sqrt[4]{c}$; **d)** $x = \sqrt[5]{d}$. Bonus points if you can also solve $x^2 = -1$. We'll get to that in Section 3.5. **E1.15** $k_e = 8.988 \times 10^9$. **E1.16 a)** $\log(2xy)$. **b)** $-\log(z)$. **c)** $\log(y)$. **d)** 3. **e)** -3. **f)** 4. **E1.17** Domain: $\mathbb{R}$. Image: $[-2, 2]$. Roots: $\{\ldots, -\frac{\pi}{2}, \frac{\pi}{2}, \frac{3\pi}{2}, \frac{5\pi}{2}, \ldots\}$. **E1.18 a)** $p(x)$ is even and has degree 4. **b)** $q(x)$ is odd and has degree 7. **E1.19 a)** $x = 5$ and $x = -3$; **b)** $x = 1 + \sqrt{3}$ and $x = 1 - \sqrt{3}$. **E1.20 a)** $(q \circ f)(x) = q(f(x)) = (x+5)^2$; $q(x)$ shifted five units to the left. **b)** $(f \circ q)(x) = x^2 + 5$; $q(x)$ shifted upward by five units. **c)** $(q \circ g)(x) = (x-6)^2$; $q(x)$ shifted six units to the right. **d)** $(q \circ h)(x) = 49x^2$; $q(x)$ horizontally compressed by a factor of seven. **E1.21** $A = 5$, $\lambda = 0.1$, and $\phi = \frac{\pi}{8}$. **E1.22** $f(x) = x^2 - 2x + 5$. **E1.23** $g(x) = 2\sqrt{x-3} - 2$. **E1.24** $x = \sqrt{21}$. **E1.25** $V = 33.51$ and $A = 50.26$. **E1.26** Length of track $= 5C = 5\pi d = 11.47$ m. **E1.27** $x = 5\cos(45°) = 3.54$, $y = 5\sin(45°) = 3.54$; $C = 10\pi$. **E1.28 a)** $\frac{\pi}{6}$ rad; **b)** $\frac{\pi}{4}$ rad; **c)** $\frac{\pi}{3}$ rad; **d)** $\frac{3\pi}{2}$ rad. **E1.29 a)** -1; **b)** 1; **c)** 0. **E1.30 a)** 0; **b)** 1; **c)** $\frac{1}{2}$; **d)** 1. **E1.31 a)** $3.16\angle 18.43°$; **b)** $2.24\angle 243.43° = 2.24\angle -116.57°$; **c)** $6\angle 270° = 6\angle -90°$; **d)** $(8.66, 5)$; **e)** $(9.66, 2.59)$; **f)** $(-5, 8.66)$. **E1.32** $y = 2$. **E1.33** $c = \sqrt{a^2 - b^2}$. **E1.34** $y = \frac{1}{4f}x^2$. **E1.35 a)** $(1, \frac{1}{2})$. **b)** $(1, 2)$. **c)** $(-2, 2)$. **E1.36** $x = 2$, $y = 3$. **E1.37** $x = 5$, $y = 6$, and $z = -3$. **E1.38** $p = 7$ and $q = 3$. **E1.39 a)** \$53 974.14; **b)** \$59 209.77; **c)** \$65 948.79. **E1.40** \$32 563.11. **E1.41 a)** $\{2, 4, 6, 7\}$; **b)** $\{1, 2, 3, 4, 5, 6\}$; **c)** $\{1, 3, 5\}$; **d)** $\varnothing$; **e)** $\{1, 2, 3, 4, 5, 6, 7\}$; **f)** $\{7\}$; **g)** $\{2, 4, 6, 7\}$; **h)** $\varnothing$. **E1.42 a)** $(-\infty, \frac{3}{2})$; **b)** $(-\infty, -5]$; **c)** $(-1, 4)$; **d)** $(4, \infty)$; **e)** $[\frac{14}{3}, \infty)$; **f)** $(-\infty, -4] \cup [2, \infty)$.

Chapter 2 solutions

Answers to problems

P2.1 The particle is gaining speed and its motion is a UAM. **P2.2** You're not running with UVM. **P2.3** (1) Speed is decreasing. (2) Speed is increasing. **P2.4** UAM before $t = t_0$ and UVM after $t = t_0$. **P2.5** $v_A = 6$[m/s], $v_B = 6$[m/s], $v_C = 8$[m/s], $v_D = 8$[m/s], and $v_E = 4$[m/s]. **P2.6** $v(t) = 9t^2 + 10t - 3$, $a(t) = 18t + 10$. **P2.7** (1) $g_{\text{jupiter}} = 24.5$[m/s^2]. (2) $y(t) = 4 - 12.25t^2$. **P2.8** (1) $v(t) = 12t + 2$[m/s], $a(t) = 12$[m/s^2]. (2) $F = 60$[N]. **P2.9** (1) $a(t) = 2$[m/s^2], $v(t) = 2t$[m/s], and $x(t) = t^2$[m]. (2) $v(4) = 8$[m/s]. (3) $v = 6$[m/s] when $x = 9$[m]. **P2.10** (1) $v(2) = 6$[m/s], $x(2) = 4$[m]. (2) For $t > 2$[s]: $v(t) = 6 + 6(t - 2)$[m/s], $x(t) = 4 + 6(t - 2) + 3(t - 2)^2$[m]. (3) When $x = 49$[m], $t = 5$[s]. (4) $v = 12$[m/s] when $x = 13$[m]. **P2.11** (1) $v_{A\to B} = 0$[m/s], $v_{C\to D} = 2$[m/s], and $v_{E\to F} = -1.5$[m/s]. (2) From 0[s] to 2[s] the squirrel is not running, from 2[s] to 6[s] it's running forward, and from 6[s] to 9[s] it's running backward. The squirrel changes direction at $t = 6$[s]. **P2.12** (1) $v_i = 6$[m/s]. (2) $x(t) = 6t - t^2$ [m]. **P2.13** (1) $x_1(t) = \frac{3}{2}t^2$[m] and $x_2(t) = \frac{1}{2}t^2 + 3t + 4$[m]. (2) $x_1 = x_2 = 24$[m] when $t = 4$[s]. **P2.14** (1) $x_i = 7$[m] and $v_i = 5$[m/s]. (2) $v(t) = 4t + 5$ [m/s] and $a(t) = 4$[m/s^2]. (3) $x(5) = 82$[m] and $v(5) = 25$[m/s]. **P2.15** $v_i = 4\sqrt{2}$[m/s] and $t_{\text{stop}} = \sqrt{2}$[s].

Solutions to problems

P2.1 The relation between velocity and time is linear and v increases with time. The rate at which the velocity increases is constant so the acceleration is constant.

P2.2 Your velocity during the first time interval (from $t = 0$[s] to $t = 2$[s]) is 1.5[m/s]. During the second time interval your velocity is 2.5[m/s], and during the third it is 3[m/s]. Your run is not a UVM.

P2.3 When the velocity and the acceleration are in opposite directions, the car's speed will decrease. When the velocity and the acceleration are in the same direction, the car's speed will increase.

P2.4 Use $F = ma$. When F is constant, a is constant and the motion is UAM. When $F = 0$, $a = 0$ and the motion is UVM.

P2.5 To find the velocity, calculate the area under the graph and add the areas for each step. Note the area is negative from D to E.

P2.6 Differentiate the function $x(t)$ to find the velocity $v(t)$. Differentiate $v(t)$ to find the acceleration $a(t)$.

P2.7 Use $v_f^2 - v_i^2 = 2a(y_f - y_i)$ to find the acceleration. Use the general equation $y(t) = y_i + v_i t + \frac{1}{2}at^2$ and plug in $y_i = 4$[m], $v_i = 0$[m/s], and $a = -24.5$[m/s^2] to find the height as a function of time.

P2.8 Differentiate the position with respect to t to find $v(t)$, then differentiate $v(t)$ to find $a(t)$. Use $F = ma$ to find the force.

P2.9 (1) Calculate $a(t)$ from Newton's 2$^{\text{nd}}$ law to obtain $a(t) = 2$[m/s^2]. Integrate with respect to time to obtain the velocity and then integrate again to find the position. (2) Plug $t = 4$[s] into $v(t)$. (3) Use the fourth equation to find the speed after the car has travelled 9[m].

P2.10 (1) Use integration to find $v(t) = \frac{3}{2}t^2$[m/s] and $x(t) = \frac{1}{2}t^3$[m], then plug in $t = 2$[s]. (2) To describe UAM starting at $t = 2$[s], use the general equation $x(t) = x_2 + v_2(t - 2) + \frac{1}{2}a(t - 2)^2$, where $x_2 = 4$[m] and $v_2 = 6$[m/s] are the initial conditions at $t = 2$[s]. (3) Solve for t in $x(t) = 49$[m]. (4) Use $v_f^2 - v_i^2 = 2a(x - x_2)$ with $v_i = 6$[m/s], $v_f = 12$[m/s], and $x_2 = 4$[m] to find the displacement during the UAM.

P5.104 Observe that $\frac{1}{2^n+n} < \frac{1}{2^n}$ for all $n \geqslant 1$. We know $\sum_n \frac{1}{2^n}$ converges since it is a geometric series with $r = \frac{1}{2}$ and $|r| < 1$. By the direct comparison test, $\sum_n \frac{1}{2^n+n}$ also converges.

P5.105 The limit of the nth term in the series is $\lim_{n\to\infty} \frac{n^2}{2n^2+1} = \frac{1}{2} \neq 0$. The series $\sum_n \frac{n^2}{2n^2+1}$ involves a limit of an infinite number of nonzero terms; therefore it must be divergent.

P5.109 In each case we use the formula for the geometric series $\sum_{n=0}^{\infty} ar^n = \frac{a}{1-r}$. See `bit.ly/1e9F52v` for the calculations.

P5.110 Use the geometric series formula $\sum_{n=0}^{\infty} ar^n = \frac{a}{1-r}$.

P5.112 This is a geometric series with $r = \frac{1}{2}$ and $a = \frac{2}{3}$, so $\frac{a}{1-r} = \frac{\frac{2}{3}}{1-\frac{1}{2}} = \frac{4}{3}$.

P5.119 Taking the approach suggested in the hint, we write the summation $\sum_{n=1}^{N} n$ by grouping 1 with N, 2 with $N-1$, 3 with $N-2$, and so on to obtain $\frac{N}{2}$ terms, each with value $N+1$. We thus find $\sum_{n=1}^{N} n = \frac{N}{2}(N+1) = \frac{N(N+1)}{2}$.

For the second part, we can use the "telescopic" nature of this series. Except for the first and last terms in this series, the negative part of each term $(n+1)^3 - n^3$ cancels the positive part of the next term. Thus only the negative part of the first term and the positive part of the last term remain: $\sum_{n=1}^{N}(n+1)^3 - n^3 = (N+1)^3 - 1$. Using basic algebra operations, we find:

$$(n+1)^3 - n^3 = n^3 + 3n^2 + 3n + 1 - n^3 = 3n^2 + 3n + 1.$$

Since we know $\sum_{n=1}^{N}(n+1)^3 - n^3 = (N+1)^3 - 1$ from the telescopic approach, we obtain the equation

$$(N+1)^3 - 1 = \sum_{n=1}^{N} 3n^2 + 3n + 1 = 3\sum_{n=1}^{N} n^2 + 3\sum_{n=1}^{N} n + \sum_{n=1}^{N} 1.$$

Using the formulas $\sum_{n=1}^{N} n = \frac{N(N+1)}{2}$ and $\sum_{n=1}^{N} 1 = N$, we rewrite the equation as

$$(N+1)^3 - 1 = 3\sum_{n=1}^{N} n^2 + 3\frac{N(N+1)}{2} + N.$$

Isolating $\sum_{n=1}^{N} n^2$ and simplifying leads us to the desired result.

P5.120 Using the definition of the integral as a the limit of a Riemann sum from page 372, we can write the following formula for the integral of $f(x)$ between $x = \alpha$ and $x = \beta$:

$$\int_{\alpha}^{\beta} f(x)\,dx \overset{\text{def}}{=} \lim_{n\to\infty} \sum_{k=1}^{n} f(a + k\Delta x)\Delta x,$$

where $\Delta x = \frac{\beta-\alpha}{n}$ denotes the width of the rectangles used to approximate the area under the curve. For the integral of the function $f(x) = mx + b$, the formula becomes

$$\int_{\alpha}^{\beta} mx + b\,dx = \lim_{n\to\infty} \sum_{k=1}^{n} (m(\alpha + k\Delta x) + b)\,\Delta x$$

$$= \lim_{n\to\infty} \Delta x \left((m\alpha + b) \sum_{k=1}^{n} 1 + m(\Delta x) \sum_{k=1}^{n} k \right)$$

$$= \lim_{n\to\infty} \Delta x \left((m\alpha + b)n + m(\Delta x)\frac{n(n+1)}{2} \right).$$

After taking the limit $n \to \infty$, and after several steps of simplification, we obtain the final answer, $\int_{\alpha}^{\beta}(mx + b)\,dx = b(\beta - \alpha) + \frac{m}{2}(\beta^2 - \alpha^2)$.

P5.121 Using the definition of the integral as an infinite Riemann sum and taking $x_k = 0 + \Delta x k$ and $\Delta x = \frac{d-0}{n}$, we obtain the following formula for the area under the curve of $f(x) = ax^2 + bx + c$ between $x = 0$ and $x = d$:

$$\int_0^d ax^2 + bx + c\,dx = \lim_{n\to\infty}\sum_{k=1}^n \left(a(\Delta x)^2 k^2 + b(\Delta x)k + c \right) \Delta x$$

$$= \lim_{n\to\infty} \Delta x \left(a(\Delta x)^2 \sum_{k=1}^n k^2 + b(\Delta x) \sum_{k=1}^n k + c \sum_{k=1}^n 1 \right)$$

$$= \lim_{n\to\infty} \Delta x \left(a(\Delta x)^2 \sum_{k=1}^n \frac{n(n+1)(2n+1)}{6} + b(\Delta x)\frac{n(n+1)}{2} + cn \right),$$

which becomes $\frac{1}{3}ad^3 + \frac{1}{2}bd^2 + cd$ after using $\Delta x = \frac{d}{n}$ and taking the limit.

P5.122 The surface area of a sphere of radius R can be computed by splitting it into narrow, circular strips of area dA, with a radius varying according to $f(x) = \sqrt{R^2 - x^2}$. The width of each strip is given by the arc length formula, $d\ell = \sqrt{1 + (f'(x))^2}\,dx$. To compute $A_{\text{sphere}} = \int_{-R}^R 2\pi f(x)\sqrt{1 + (f'(x))^2}\,dx$, we first find $f'(x) = \frac{-x}{\sqrt{R^2-x^2}}$, then obtain $\sqrt{1 + (f'(x))^2} = \frac{R}{\sqrt{R^2-x^2}}$. Substituting into the formula, we find $A_{\text{sphere}} = \int_{-R}^R 2\pi\sqrt{R^2 - x^2}\frac{R}{\sqrt{R^2-x^2}}\,dx$, which simplifies to $2\pi R \int_{-R}^R 1\,dx$, which gives $4\pi R^2$ as expected.

P5.125 Using the disk method, we find the volume of Torricelli's trumpet between $x = 1$ and $x = a$ is $\int_1^a \pi(f(x))^2\,dx = \pi\int_1^a \frac{1}{x^2}\,dx = \pi\left(1 - \frac{1}{a}\right)$. To find the total volume, we compute the limit $\lim_{a\to\infty} \pi\left(1 - \frac{1}{a}\right) = \pi$.

P5.126 To evaluate $I_{\text{sph.shell}} = \int_{-R}^R (f(x))^2\,\sigma 2\pi f(x)\sqrt{1 + (f'(x))^2}\,dx$, we first compute $f'(x) = \frac{-x}{\sqrt{R^2-x^2}}$ then obtain $\sqrt{1 + (f'(x))^2} = \frac{R}{\sqrt{R^2-x^2}}$. Substituting these calculations into the integral, we find $I_{\text{sph.shell}} = 2\pi\sigma R \int_{-R}^R (R^2 - x^2)\,dx$, which is the integral of a polynomial.

P5.127 The formula for computing the volume of a sphere using the spherical shell method is $V_{\text{sphere}} = \int_0^R 2\pi r h(r)\,dr$, where $h(r) = 2\sqrt{R^2 - r^2}$ is the height of the thin cylindrical shell at radius r, and $2\pi r$ is its circumference. To evaluate the integral $I_{\text{sphere}} = \int r^2 dm$, we define $dm = \rho A(r)dr = \rho 2\pi r h(r)\,dr$ and scale the contributions of each dm by an extra factor r^2. We thus obtain $I_{\text{sphere}} = \int_0^R r^2 \rho 2\pi r 2\sqrt{R^2 - r^2}\,dr$. Computing the integral $\int_0^R r^3\sqrt{R^2 - r^2}\,dr$ requires several steps, and I encourage you to perform the calculation on your own. The substitutions $r = R\sin\theta$ and later $u = \cos\theta$ might come in handy.

P5.128 The integral to calculate, $\int_{-R}^R I_{\text{disk}}(x)\,dx = \int_{-R}^R \frac{1}{2}m(x)[r(x)]^2\,dx$, becomes $\frac{1}{2}\pi\rho \int_{-R}^R (f(x))^4\,dx$ after substituting all the known values. After expanding $(f(x))^4$, we find $I_{\text{sphere}} = \frac{1}{2}\pi\rho \int_{-R}^R (R^4 - 2R^2 x^2 + x^4)\,dx$, which is straightforward to compute.

Appendix B

Notation

This appendix contains a summary of the notation used in this book.

Math notation

Expression	Read as	Used to denote
a, b, x, y		variables
$=$	is equal to	expressions that have the same value
$\overset{\text{def}}{=}$	is defined as	a new variable definition
$a + b$	a plus b	the combined lengths of a and b
$a - b$	a minus b	the difference in lengths between a and b
$a \times b = ab$	a times b	the area of a rectangle
$a^2 = aa$	a squared	the area of a square of side length a
$a^3 = aaa$	a cubed	the volume of a cube of side length a
a^n	a to the n	a multiplied by itself n times
$\sqrt{a} = a^{\frac{1}{2}}$	square root of a	the side length of a square of area a
$\sqrt[3]{a} = a^{\frac{1}{3}}$	cube root of a	the side length of a cube with volume a
$a/b = \frac{a}{b}$	a divided by b	a parts of a whole split into b parts
$a^{-1} = \frac{1}{a}$	one over a	division by a
$f(x)$	f of x	the function f applied to input x
f^{-1}	f inverse	the inverse function of $f(x)$
$f \circ g$	f compose g	function composition; $f \circ g(x) = f(g(x))$
e^x	e to the x	the exponential function base e
$\ln(x)$	natural log of x	the logarithm base e
a^x	a to the x	the exponential function base a
$\log_a(x)$	log base a of x	the logarithm base a
θ, ϕ	*theta*, *phi*	angles
$\sin, \cos, \tan$	sin, cos, tan	trigonometric ratios
$\%$	percent	proportions of a total; $a\% = \frac{a}{100}$

Set notation

You don't need a lot of fancy notation to understand mathematics. It really helps, though, if you know a little bit of set notation.

Symbol	Read as	Denotes
$\{\dots\}$	the set ...	define a sets
$\|$	such that	describe or restrict the elements of a set
$\mathbb{N}$	the naturals	the set $\mathbb{N} \stackrel{\text{def}}{=} \{0, 1, 2, \dots\}$. Note $\mathbb{N}^* \stackrel{\text{def}}{=} \mathbb{N}\backslash\{0\}$.
$\mathbb{Z}$	the integers	the set $\mathbb{Z} \stackrel{\text{def}}{=} \{\dots, -2, -1, 0, 1, 2, 3, \dots\}$
$\mathbb{Q}$	the rationals	the set of fractions of integers
$\mathbb{R}$	the reals	the set of real numbers
$\mathbb{C}$		the set of complex numbers
$\subset$	subset	one set strictly contained in another
$\subseteq$	subset or equal	containment or equality
$\cup$	union	the combined elements from two sets
$\cap$	intersection	the elements two sets have in common
$S\backslash T$	S set minus T	the elements of S that are not in T
$a \in S$	a in S	a is an element of set S
$a \notin S$	a not in S	a is not an element of set S
$\forall x$	for all x	a statement that holds for all x
$\exists x$	there exists x	an existence statement
$\nexists x$	there doesn't exist x	a non-existence statement

An example of a condensed math statement that uses set notation is "$\nexists m, n \in \mathbb{Z}$ such that $\frac{m}{n} = \sqrt{2}$," which reads "there don't exist integers m and n whose ratio equals $\sqrt{2}$." Since we identify the set of fractions of integers with the rationals, this statement is equivalent to the shorter "$\sqrt{2} \notin \mathbb{Q}$," which reads "$\sqrt{2}$ is irrational."

Complex numbers notation

Expression	Denotes
$\mathbb{C}$	the set of complex numbers $\mathbb{C} \stackrel{\text{def}}{=} \{a + bi \mid a, b \in \mathbb{R}\}$
i	the unit imaginary number $i \stackrel{\text{def}}{=} \sqrt{-1}$ or $i^2 = -1$
$\mathrm{Re}\{z\} = a$	real part of $z = a + bi$
$\mathrm{Im}\{z\} = b$	imaginary part of $z = a + bi$
$|z|\angle\varphi_z$	polar representation of $z = |z|\cos\varphi_z + i|z|\sin\varphi_z$
$|z| = \sqrt{a^2 + b^2}$	magnitude of $z = a + bi$
$\varphi_z = \tan^{-1}(b/a)$	phase or argument of $z = a + bi$
$\bar{z} = a - bi$	complex conjugate of $z = a + bi$

Vectors notation

Expression	Denotes
$\mathbb{R}^n$	the set of n-dimensional vectors
$\vec{v}$	a vector
(v_x, v_y)	vector in component notation
$v_x\hat{\imath} + v_y\hat{\jmath}$	vector in unit vector notation
$\|\vec{v}\|\angle\theta$	vector in length-and-direction notation
$\|\vec{v}\|$	length of the vector $\vec{v}$
θ	angle the vector $\vec{v}$ makes with the x-axis
$\hat{v} \overset{\text{def}}{=} \dfrac{\vec{v}}{\|\vec{v}\|}$	unit vector in the same direction as $\vec{v}$
$\vec{u} \cdot \vec{v}$	dot product of the vectors $\vec{u}$ and $\vec{v}$
$\vec{u} \times \vec{v}$	cross product of the vectors $\vec{u}$ and $\vec{v}$

Mechanics notation

Expression	Denotes
$x(t)$	position of an object as a function of time
$v(t)$	velocity of an object as a function of time
$a(t)$	acceleration of an object as a function of time
m	mass of an object
$\vec{F}$	a force
$\vec{N}$	normal force
$\vec{F}_{fs}$	static force of friction
$\vec{F}_{fk}$	kinetic force of friction
$\vec{F}_g = \vec{W}$	gravitational force; the weight of an object
U_g	gravitational potential energy
$\vec{F}_s$	force of a spring
U_s	spring potential energy
$\vec{p}$	momentum of a moving object
K	kinetic energy
$\theta(t)$	angular position of a rotating object over time
$\omega(t)$	angular velocity of an object as a function of time
$\alpha(t)$	angular acceleration of an object as a function of time
I_{obj}	moment of inertia of an object
$\mathcal{T}$	torque
L	angular momentum of a spinning object
K_r	rotational kinetic energy of a spinning object

Calculus notation

Expression	Denotes
∞	infinity
ϵ, δ	the Greek letters *epsilon* and *delta*
$f(x)$	a function of the form $f : \mathbb{R} \to \mathbb{R}$
$\lim\limits_{x \to \infty} f(x)$	limit of $f(x)$ as x goes to infinity
$\lim\limits_{x \to a^+} f(x)$	limit of $f(x)$ as x approaches a from the right
$\lim\limits_{x \to a^-} f(x)$	limit of $f(x)$ as x approaches a from the left
$\lim\limits_{x \to a} f(x)$	limit of $f(x)$ as x goes to a
$f'(x)$	derivative of $f(x)$
$f''(x)$	second derivative of $f(x)$
$\frac{d}{dx}$	derivative operator
$F(x)$	antiderivative function of $f(x)$
$\int f(x)\, dx$	indefinite integral of $f(x)$ (spoiler: $\int f(x)\, dx = F(x) + C$)
$\int_a^b f(x)\, dx$	definite integral of $f(x)$ between $x = a$ and $x = b$
$F(x)\big|_\alpha^\beta$	change in $F(x)$ between α and β: $F(x)\big|_\alpha^\beta = F(\beta) - F(\alpha)$
$\sum\limits_{k=1}^{N} s_k$	summation of N terms $s_1 + s_2 + \cdots + s_N$
a_n	sequence $a_n : \mathbb{N} \to \mathbb{R}$, also denoted $(a_0, a_1, a_2, a_3, \dots)$
$\sum\limits_{n=0}^{\infty} a_n$	series $\sum a_n$, which is the infinite sum of the sequence a_n

Appendix C

Constants, units, and conversion ratios

In this appendix you will find a number of tables of useful information that you might need when solving math and physics problems.

Fundamental constants of Nature

Many of the equations of physics include constants as parameters of the equation. For example, Newton's law of gravitation says that the force of gravity between two objects of mass M and m separated by a distance r is $F_g = \frac{GMm}{r^2}$, where G is Newton's gravitational constant.

Symbol	Value	Units	Name
G	6.67384×10^{-11}	$\mathrm{m^3\,kg^{-1}s^{-2}}$	gravitational constant
g	$9.80665 \approx 9.81$	$\mathrm{m\,s^{-2}}$	Earth free-fall acceleration
m_P	1.672621×10^{-27}	kg	proton mass
m_e	9.109382×10^{-31}	kg	electron mass
N_A	$6.02214076 \times 10^{23}$	$\mathrm{mol^{-1}}$	Avogadro's number
k_B	1.380648×10^{-23}	$\mathrm{J\,K^{-1}}$	Boltzmann's constant
R	8.3144621	$\mathrm{J\,K^{-1}\,mol^{-1}}$	gas constant $R = N_A k_B$
μ_0	1.256637×10^{-6}	$\mathrm{N\,A^{-2}}$	permeability of free space
ε_0	8.854187×10^{-12}	$\mathrm{F\,m^{-1}}$	permittivity of free space
c	$299\,792\,458$	$\mathrm{m\,s^{-1}}$	speed of light $c = \frac{1}{\sqrt{\mu_0 \varepsilon_0}}$
e	1.602176×10^{-19}	C	elementary charge
h	6.626069×10^{-34}	$\mathrm{J\,s}$	Planck's constant

Units

The International System of Units (*Système International*) defines seven base units for measuring physical quantities.

Name	Sym.	Measures	Definition
metre	m	length	The distance travelled by light in vacuum during $\frac{1}{299792458}$th of a second.
second	s	time	The time for 9192631770 transitions in the ground state of the caesium-133 atom.
kilogram	kg	mass	Defined in terms of the speed of light c, an atomic transition frequency of the caesium-133 atom, and Planck's constant h.
Ampere	A	electric current	One ampere is the current that has to flow in two infinitely long wires placed a distance 1[m] apart, to produce a force between them of 2×10^{-7}[N/m].
Kelvin	K	temperature	The Kelvin is $\frac{1}{273.16}$ of the thermodynamic temperature of the triple point of water.
mole	mol	# of atoms	One mole is how many carbon atoms are in 0.012[kg] of carbon-12.
candela	cd	light intensity	One candela is defined as the luminous intensity of a monochromatic source with a particular frequency and radiant intensity.

Derived units

The base SI units cover most of the fundamental quantities. Other physical units are defined as combinations of the basic units.

Name	Sym.	Measures	Definition	SI equivalent
Hertz	Hz	frequency		s^{-1}
Newton	N	force		$kg\,m\,s^{-2}$
Pascal	Pa	pressure	N/m^2	$kg\,m^{-1}\,s^{-2}$
Joule	J	energy, work, heat	$N\,m$	$kg\,m^2\,s^{-2}$
Watt	W	power	J/s	$kg\,m^2\,s^{-3}$
Coulomb	C	electric charge		$s\,A$
Volt	V	voltage, electric potential		$kg\,m^2\,s^{-3}\,A^{-1}$
Ohm	Ω	resistance, reactance	V/A	$kg\,m^2\,s^{-3}\,A^{-2}$
Siemens	S	electrical conductance	A/V	$kg^{-1}\,m^{-2}\,s^3\,A^2$
Farad	F	capacitance	C/V	$kg^{-1}\,m^{-2}\,s^4\,A^2$
Tesla	T	magnetic field strength		$kg\,s^{-2}\,A^{-1}$
Henry	H	inductance	$\Omega\,s$	$kg\,m^2\,s^{-2}\,A^{-2}$
Weber	Wb	magnetic flux	$T\,m^2$	$kg\,m^2\,s^{-2}\,A^{-1}$

Other units and conversions

We often measure physical quantities like length, weight, and velocity in nonstandard units like feet, pounds, and miles per hour. The following table lists the conversion ratios which are required to covert these nonstandard measurement units to SI units.

Dimension	Symb.	Name	Conversion
length	Å	Angstrom	$1[\text{Å}] = 10^{-10}[\text{m}] = 0.1[\text{nm}]$
	in, "	inch	$1[\text{in}] = 2.54[\text{cm}]$
	ft	foot	$1[\text{ft}] = 12[\text{in}] = 0.3048[\text{m}]$
	yd	yard	$1[\text{yd}] = 3[\text{ft}] = 0.9144[\text{m}]$
	mi	mile	$1[\text{mi}] = 5280[\text{ft}] = 1609.344[\text{m}]$
	nmi	nautical mile	$1[\text{nmi}] = 1852[\text{m}]$
	ly	light-year	$1[\text{ly}] = 9.460\,730\,472 \times 10^{15}[\text{m}]$
area	in^2	square inch	$1[\text{in}^2] = 6.452 \times 10^{-4}[\text{m}^2]$
	ft^2	square foot	$1[\text{ft}^2] = 9.290 \times 10^{-2}[\text{m}^2]$
			$1[\text{m}^2] = 110.764[\text{ft}^2]$
	ac	acre	$1[\text{ac}] = 4840[\text{yd}^2] = 4046.856[\text{m}^2]$
	ha	hectare	$1[\text{ha}] = 10\,000[\text{m}^2]$
	mi^2	square mile	$1[\text{mi}^2] = 2.589\,988 \times 10^6[\text{m}^2]$
volume	L	litre	$1[\text{L}] = 1[\text{dm}^3] = \frac{1}{1000}[\text{m}^3]$
	gal(US)	gallon (fluid)	$1[\text{gal}] = 3.785[\text{L}]$
weight	lb	pound	$1[\text{lb}] = 0.454[\text{kg}] = 453.592[\text{g}]$
	t	tonne	$1[\text{t}] = 1000[\text{kg}]$
angle	rad	radian	$1[\text{turn}] = 2\pi[\text{rad}]$
	°	degree	$360[°] = 2\pi[\text{rad}]$
	rev	revolution	$1[\text{rev}] = 360[°] = 2\pi[\text{rad}]$
	grad	gradian	$1[\text{grad}] = \frac{1}{400}[\text{rev}] = 0.9[°]$
time	min	minute	$1[\text{min}] = 60[\text{s}]$
	h	hour	$1[\text{h}] = 60[\text{min}] = 3600[\text{s}]$
velocity	km/h	km per hour	$1[\text{km/h}] = \frac{1}{3.6}[\text{m/s}] = 0.2\overline{7}[\text{m/s}]$
	mph	mile per hour	$1[\text{mph}] = 0.447[\text{m/s}] = 1.61[\text{km/h}]$
temperature	°C	Celsius	$x[°\text{C}] = (x + 273.15)[°\text{K}]$
	°F	Fahrenheit	$x[°\text{F}] = \frac{5}{9}(x + 459.67)[°\text{K}]$
			$x[°\text{F}] = \frac{5}{9}(x - 32)[°\text{C}]$
pressure	atm	atmosphere	$1[\text{atm}] = 101\,325[\text{Pa}]$
	bar	bar	$1[\text{bar}] = 10^5[\text{Pa}]$

Appendix D

SymPy tutorial

Computers can be very useful for dealing with complicated math expressions or when slogging through tedious calculations. Throughout this book we used `SymPy` to illustrate several concepts from math and physics. We'll now review all the math and physics tools available through the `SymPy` command line. Don't worry if you're not a computer person; we'll only discuss concepts we covered in the book, and the computer commands we'll learn are very similar to the math operations you're already familiar with. This section also serves as a final review of the material covered in the book.

Introduction

You can use a computer algebra system (CAS) to compute complicated math expressions, solve equations, perform calculus procedures, and simulate physics systems.

All computer algebra systems offer essentially the same functionality, so it doesn't matter which system you use: there are free systems like `SymPy`, `Magma`, or `Octave`, and commercial systems like `Maple`, `MATLAB`, and `Mathematica`. This tutorial is an introduction to `SymPy`, which is a *symbolic* computer algebra system written in the programming language Python. In a symbolic CAS, numbers and operations are represented symbolically, so the answers obtained are exact. For example, the number $\sqrt{2}$ is represented in `SymPy` as the object `Pow(2,1/2)`, whereas in *numerical* computer algebra systems like `Octave`, the number $\sqrt{2}$ is represented as the approximation 1.41421356237310 (a `float`). For most purposes the approximation is okay, but sometimes approximations can lead to problems: `float(sqrt(2))*float(sqrt(2))` = 2.00000000000000044 $\neq$ 2. Because `SymPy` uses exact representations, you'll never run into such problems: `Pow(2,1/2)*Pow(2,1/2)` = 2.

This tutorial presents many explanations as code snippets. Be sure to try the code examples on your own by typing the commands into `SymPy`. It's always important to verify for yourself!

Using SymPy

The easiest way to use `SymPy`, provided you're connected to the internet, is to visit `http://live.sympy.org`. You'll be presented with an interactive prompt into which you can enter your commands—right in your browser.

If you want to use `SymPy` on your own computer, you must first install Python and the Python package `sympy`. You can then open a command prompt and start a Python session using:

```
you@host> python
Python X.Y.Z
[GCC a.b.c (Build Info)] on platform
Type "help", "copyright", or "license" for more information.
>>>
```

The `>>>` prompt indicates you're in the Python shell which accepts Python commands. Type the following in the Python shell:

```
>>> from sympy import *
>>>
```

The command `from sympy import *` imports all the `SymPy` functions into the current namespace. All `SymPy` functions are now available to you. To exit the python shell press `CTRL+D`.

For an even better experience, you can try `jupyter notebook`, which is a web interface for accessing the Python shell. Search the web for "jupyter notebook" and follow the installation instructions specific to your operating system. It's totally worth it!

Each section in this appendix begins with a python `import` statement for the functions used in that section. If you use the statement `from sympy import *` in the beginning of your code, you don't need to run these individual import statements, but I've included them so you'll know which `SymPy` vocabulary is covered in each section.

Fundamentals of mathematics

Let's begin by learning about the basic `SymPy` objects and the operations we can carry out on them. We'll learn the `SymPy` equivalents of many math verbs like: "to solve" (an equation), "to expand" (an expression), "to factor" (a polynomial).

Numbers

```
>>> from sympy import  sympify, S, evalf, N
```
In Python, there are two types of number objects: `ints` and `floats`.
```
>>> 3
3                        # an int
>>> 3.0
3.0                      # a float
```
Integer objects in Python are a faithful representation of the set of integers $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$. Floating point numbers are approximate representations of the reals $\mathbb{R}$. A floating point number has 16 decimals of precision.

Special care is required when specifying rational numbers if you want to get exact answers. If you try to divide two numbers, Python will compute a floating point approximation:
```
>>> 1/7
0.14285714285714285      # a float
```
The floating point number 0.14285714285714285 is an approximation of the exact number $\frac{1}{7} \in \mathbb{Q}$. The `float` approximation has 16 decimals while the decimal expansion of $\frac{1}{7}$ is infinitely long. To obtain an *exact* representation of $\frac{1}{7}$ you need to create a SymPy expression. You can `sympify` any expression using the shortcut function `S()`:
```
>>> S('1/7')             # = sympify('1/7')
1/7                      # = Rational(1,7)
```
Note the input to `S()` is specified as a text string delimited by quotes. We could have achieved the same result using `S('1')/7` since a SymPy object divided by an `int` is a SymPy object.

Except for the tricky Python division operator, other math operators like addition +, subtraction -, and multiplication * work as you would expect. The syntax ** is used to denote exponentiation:
```
>>> 2**10                # same as S('2^10')
1024
```
When solving math problems, it's best to work with SymPy objects, and wait to compute the numeric answer in the end. To obtain a numeric approximation of a SymPy object as a `float`, call its `.evalf()` method:
```
>>> pi
pi
>>> pi.evalf()           # = pi.n() = N(pi)
3.14159265358979
```
The method `.n()` is equivalent to `.evalf()`. The global SymPy function `N()` can also be used to to compute numerical values. You can easily change the number of digits of precision of the approximation. Enter `pi.n(400)` to obtain an approximation of π to 400 decimals.

Symbols

```
>>> from sympy import Symbol, symbols
```

Python is a civilized language so there's no need to define variables before assigning values to them. When you write `a = 3`, you define a new name `a` and set it to the value 3. You can now use the name `a` in subsequent calculations.

Most interesting SymPy calculations require us to define `symbols`, which are the SymPy objects for representing variables and unknowns. For your convenience, when `live.sympy.org` starts, it runs the following commands automatically:

```
>>> from sympy import *
>>> x, y, z, t = symbols('x y z t')
>>> k, m, n = symbols('k m n', integer=True)
>>> f, g, h = symbols('f g h', cls=Function)
```

The first statement imports all the SymPy functions. The other three statements define some generic symbols x, y, z, and t, and several other symbols with special properties.

Note the difference between the following two statements:

```
>>> x + 2
x + 2                   # an Add expression
>>> p + 2
NameError: name 'p' is not defined
```

The name x is defined as a symbol, so SymPy knows that x + 2 is an expression; but the variable p is not defined, so SymPy doesn't know what to make of p + 2. To use p in expressions, you must first define it as a symbol:

```
>>> p = Symbol('p')    # the same as p = symbols('p')
>>> p + 2
p + 2                   # = Add(Symbol('p'), Integer(2))
```

You can define a sequence of variables using the following notation:

```
>>> a0, a1, a2, a3 = symbols('a0:4')
```

You can use any name you want for a variable, but it's best if you avoid the letters Q,C,O,S,I,N and E because they have special uses in SymPy: I is the unit imaginary number $i \stackrel{\text{def}}{=} \sqrt{-1}$, E is the base of the natural logarithm, S() is the `sympify` function, N() is used to obtain numeric approximations, and O is used for big-O notation.

Expressions

```
>>> from sympy import simplify, factor, expand, collect
```

You define SymPy expressions by combining symbols with basic math operations and other functions:

```
>>> expr = 2*x + 3*x - sin(x) - 3*x + 42
>>> simplify(expr)                # simplify the expression
2*x - sin(x) + 42
```

The function `simplify` can be used on any expression to simplify it. The examples below illustrate other useful `SymPy` functions that correspond to common mathematical operations on expressions:

```
>>> factor( x**2-2*x-8 )          # factor a polynomial
(x - 4)*(x + 2)
>>> expand( (x-4)*(x+2) )         # expand and expression
x**2 - 2*x - 8
>>> collect(x**2+x*b+a*x+a*b, x)   # collect like terms in x
x**2 + (a+b)*x + a*b
```

To substitute a given value into an expression, call the `.subs()` method, passing in a python dictionary object `{ key:val, ... }` with the symbol–value substitutions you want to make:

```
>>> expr  = sin(x) + cos(y)       # define an expression
>>> expr
sin(x) + cos(y)
>>> expr.subs({x:1, y:2})          # subs. x=1,y=1 in expr
sin(1) + cos(2)
>>> expr.subs({x:1, y:2}).n()      # compute numeric value
0.425324148260754
```

Note how we used `.n()` to obtain the expression's numeric value.

Solving equations

```
>>> from sympy import solve
```

The function `solve` is the main workhorse in `SymPy`. This incredibly powerful function knows how to solve all kinds of equations. In fact `solve` can solve pretty much *any* equation! When high school students learn about this function, they get really angry—why did they spend five years of their life learning to solve various equations by hand, when all along there was this `solve` thing that could do all the math for them? Don't worry, learning math is *never* a waste of time.

The function `solve` takes two arguments. Use `solve(expr,var)` to solve the equation `expr==0` for the variable `var`. You can rewrite any equation in the form `expr==0` by moving all the terms to one side of the equation; the solutions to $A(x) = B(x)$ are the same as the solutions to $A(x) - B(x) = 0$.

For example, to solve the quadratic equation $x^2 + 2x - 8 = 0$, use

```
>>> solve( x**2 + 2*x - 8, x)
[2, -4]
```

In this case the equation has two solutions so `solve` returns a list. Check that $x = 2$ and $x = -4$ satisfy the equation $x^2 + 2x - 8 = 0$.

The best part about `solve` and SymPy is that you can obtain symbolic answers when solving equations. Instead of solving one specific quadratic equation, we can solve all possible equations of the form $ax^2 + bx + c = 0$ using the following steps:

```
>>> a, b, c = symbols('a b c')
>>> solve( a*x**2 + b*x + c, x)
[(-b+sqrt(b**2 - 4*a*c))/(2*a), (-b-sqrt(b**2-4*a*c))/(2*a)]
```

In this case `solve` calculated the solution in terms of the symbols a, b, and c. You should be able to recognize the expressions in the solution—it's the quadratic formula $x_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$.

To solve a *system of equations*, you can feed `solve` with the list of equations as the first argument, and specify the list of unknowns you want to solve for as the second argument. For example, to solve for x and y in the system of equations $x + y = 3$ and $3x - 2y = 0$, use

```
>>> solve([x + y - 3, 3*x - 2*y], [x, y])
{x: 6/5, y: 9/5}
```

The function `solve` is like a Swiss Army knife you can use to solve all kind of problems. Suppose you want to *complete the square* in the expression $x^2 - 4x + 7$, that is, you want to find constants h and k such that $x^2 - 4x + 7 = (x - h)^2 + k$. There is no special "complete the square" function in SymPy, but you can call `solve` on the equation $(x - h)^2 + k - (x^2 - 4x + 7) = 0$ to find the unknowns h and k:

```
>>> h, k = symbols('h k')
>>> solve( (x-h)**2 + k  - (x**2-4*x+7), [h,k] )
[(2, 3)]                                  # so h = 2 and k = 3
>>> expand( (x-2)**2+3 )                  # verify...
x**2 - 4*x + 7
```

Learn the basic SymPy commands and you'll never need to suffer another tedious arithmetic calculation painstakingly performed by hand again!

Rational functions

```
>>> from sympy import together, apart
```

By default, SymPy will not combine or split rational expressions. You need to use `together` to symbolically calculate the addition of fractions:

```
>>> a, b, c, d = symbols('a b c d')
>>> a/b + c/d
a/b + c/d
>>> together(a/b + c/d)
(a*d + b*c)/(b*d)
```

Alternately, if you have a rational expression and want to divide the numerator by the denominator, use the `apart` function:

```
>>> apart( (x**2+x+4)/(x+2)  )
x - 1  +  6/(x + 2)
```

Exponentials and logarithms

Euler's number $e = 2.71828\ldots$ is defined one of several ways,

$$e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = \lim_{\epsilon\to 0}(1+\epsilon)^{1/\epsilon} = \sum_{n=0}^{\infty}\frac{1}{n!},$$

and is denoted E in SymPy. Using `exp(x)` is equivalent to E**x.

The functions `log` and `ln` both compute the logarithm base e:

```
>>> log(E**3)    # same as ln(E**3)
3
```

By default, SymPy assumes the inputs to functions like `exp` and `log` are complex numbers, so it will not expand certain logarithmic expressions. However, indicating to SymPy that the inputs are positive real numbers will make the expansions work:

```
>>> x, y = symbols('x y')
>>> expand( log(x*y) )
log(x*y)
>>> a, b = symbols('a b', positive=True)
>>> expand( log(a*b) )
log(a) + log(b)
```

Polynomials

Let's define a polynomial P with roots at $x = 1$, $x = 2$, and $x = 3$:

```
>>> P = (x-1)*(x-2)*(x-3)
>>> P
(x - 1)*(x - 2)*(x - 3)
```

To see the expanded version of the polynomial, call its `expand` method:

```
>>> expand(P)
x**3 - 6*x**2 + 11*x - 6
```

When the polynomial is expressed in it's expanded form $P(x) = x^3 - 6^2 + 11x - 6$, we can't immediately identify its roots. This is why the factored form $P(x) = (x - 1)(x - 2)(x - 3)$ is preferable. To factor a polynomial, call its `factor` method or `simplify` it:

```
>>> factor(P)
(x - 1)*(x - 2)*(x - 3)
>>> simplify(P)
(x - 1)*(x - 2)*(x - 3)
```

Recall that the roots of the polynomial $P(x)$ are defined as the solutions to the equation $P(x) = 0$. We can use the `solve` function to find the roots of the polynomial:

```
>>> roots = solve(P,x)
>>> roots
[1, 2, 3]
# let's check if P equals (x-1)(x-2)(x-3)
>>> simplify( P  -  (x-roots[0])*(x-roots[1])*(x-roots[2]) )
0
```

Equality checking

In the last example, we used the `simplify` function on the difference of two expressions to check whether they were equal. This way of checking equality works because $P = Q$ if and only if $P - Q = 0$. To know whether $P = Q$, we can calculate `simplify(P-Q)` and see if the result equals 0. This is the best way to check whether two expressions are equal in SymPy because it attempts all possible simplifications when comparing the expressions. Below is a list of other ways to check whether two quantities are equal, with example cases where equality fails to be detected:

```
>>> P = (x-5)*(x+5)
>>> Q = x**2 - 25
>>> P == Q                                  # fail
False
>>> P - Q == 0                              # fail
False
>>> simplify(P - Q)                         # works!
0
>>> sin(x)**2 + cos(x)**2   == 1            # fail
False
>>> simplify( sin(x)**2 + cos(x)**2 - 1 )   # works!
0
```

Trigonometry

```
from sympy import sin, cos, tan, trigsimp, expand_trig
```

The trigonometric functions `sin` and `cos` take inputs in radians:

```
>>> sin(pi/6)
1/2
>>> cos(pi/6)
sqrt(3)/2
```

For angles in degrees, you need a conversion factor of $\frac{\pi}{180}[\text{rad}/^\circ]$:

```
>>> sin(30*pi/180)                # 30 deg = pi/6 rads
1/2
```

The inverse trigonometric functions $\sin^{-1}(x) = \arcsin(x)$ and $\cos^{-1}(x) = \arccos(x)$ are used as follows:

```
>>> asin(1/2)
pi/6
>>> acos(sqrt(3)/2)
pi/6
```

Recall that $\tan(x) = \frac{\sin(x)}{\cos(x)}$. The inverse function of $\tan(x)$ is $\tan^{-1}(x) = \arctan(x) = \texttt{atan(x)}$

```
>>> tan(pi/6)
1/sqrt(3)                         # = ( 1/2 )/( sqrt(3)/2 )
>>> atan( 1/sqrt(3) )
pi/6
```

The function `acos` returns angles in the range $[0, \pi]$, while `asin` and `atan` return angles in the range $[-\frac{\pi}{2}, \frac{\pi}{2}]$.

Here are some trigonometric identities that SymPy knows:

```
>>> sin(x) == cos(x - pi/2)
True
>>> simplify( sin(x)*cos(y)+cos(x)*sin(y) )
sin(x + y)
>>> e = 2*sin(x)**2 + 2*cos(x)**2
>>> trigsimp(e)
2
>>> trigsimp(log(e))
log(2*sin(x)**2 + 2*cos(x)**2)
>>> trigsimp(log(e), deep=True)
log(2)
>>> simplify( sin(x)**4 - 2*cos(x)**2*sin(x)**2 + cos(x)**4 )
cos(4*x)/2 + 1/2
```

The function `trigsimp` does essentially the same job as `simplify`.

If instead of simplifying you want to expand a trig expression, you should use `expand_trig`, because the default `expand` won't touch trig functions:

```
>>> expand(sin(2*x))
sin(2*x)
>>> expand_trig(sin(2*x))    # = expand(sin(2*x), trig=True)
2*sin(x)*cos(x)
```

Complex numbers

```
>>> from sympy import I, re, im, Abs, arg, conjugate
```
Consider the quadratic equation $x^2 = -1$. There are no real solutions to this equation, but we can define an imaginary number $i = \sqrt{-1}$ (denoted I in SymPy) that satisfies this equation:
```
>>> I*I
-1
>>> solve( x**2 + 1 , x)
[I, -I]
```
The solutions are $x = i$ and $x = -i$, and indeed we can verify that $i^2 + 1 = 0$ and $(-i)^2 + 1 = 0$ since $i^2 = -1$.

The complex numbers $\mathbb{C}$ are defined as $\{a + bi \mid a, b \in \mathbb{R}\}$. Complex numbers contain a real part and an imaginary part:
```
>>> z = 4 + 3*I
>>> z
4 + 3*I
>>> re(z)
4
>>> im(z)
3
```
The *polar* representation of a complex number is $z = |z| \angle \theta = |z| e^{i\theta}$. For a complex number $z = a + bi$, the quantity $|z| = \sqrt{a^2 + b^2}$ is known as the *absolute value* of z, and θ is its *phase* or its *argument*:
```
>>> Abs(z)
5
>>> arg(z)
atan(3/4)
```
The complex conjugate of $z = a + bi$ is the number $\bar{z} = a - bi$, which has the same absolute value as z but opposite phase:
```
>>> conjugate( z )
4 - 3*I
```
Complex conjugation is important for computing the absolute value of z ($|z| = \sqrt{z\bar{z}}$) and for division by z ($\frac{1}{z} = \frac{\bar{z}}{|z|^2}$).

Euler's formula

```
>>> from sympy import expand, rewrite
```
Euler's formula shows an important relation between the exponential function e^x and the trigonometric functions $\sin(x)$ and $\cos(x)$:

$$e^{ix} = \cos x + i \sin x.$$

To obtain this result in SymPy, you must specify that the number x is real and also tell expand that you're interested in complex expansions:

```
>>> x = symbols('x', real=True)
>>> expand(exp(I*x), complex=True)
cos(x) + I*sin(x)
>>> re( exp(I*x) )
cos(x)
>>> im( exp(I*x) )
sin(x)
```

Basically, $\cos(x)$ is the real part of e^{ix}, and $\sin(x)$ is the imaginary part of e^{ix}. Whaaat? I know it's weird, but weird things are bound to happen when you input imaginary numbers to functions.

Calculus

Calculus is the study of the properties of functions. The operations of calculus are used to describe the limit behaviour of functions, calculate their rates of change, and calculate the areas under their graphs. In this section we'll learn about the SymPy functions for calculating limits, derivatives, integrals, and summations.

Infinity

```
from sympy import oo
```

The infinity symbol is denoted oo (two lowercase os) in SymPy. Infinity is not a number but a process: the process of counting forever. Thus, $\infty + 1 = \infty$, ∞ is greater than any finite number, and $1/\infty$ is an infinitely small number. Sympy knows how to correctly treat infinity in expressions:

```
>>> oo+1
oo
>>> 5000 < oo
True
>>> 1/oo
0
```

Limits

```
from sympy import limit
```

We use limits to describe, with mathematical precision, infinitely large quantities, infinitely small quantities, and procedures with infinitely many steps.

The number e is defined as the limit $e = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n$:

```
>>> limit( (1+1/n)**n, n, oo)
E          # = 2.71828182845905
```

This limit expression describes the annual growth rate of a loan with a nominal interest rate of 100% and infinitely frequent compounding. Borrow $1000 in such a scheme, and you'll owe $2718.28 after one year.

Limits are also useful to describe the behaviour of functions. Consider the function $f(x) = \frac{1}{x}$. The `limit` command shows us what happens to $f(x)$ near $x = 0$ and as x goes to infinity:

```
>>> limit( 1/x, x, 0, dir="+")
oo
>>> limit( 1/x, x, 0, dir="-")
-oo
>>> limit( 1/x, x, oo)
0
```

As x becomes larger and larger, the fraction $\frac{1}{x}$ becomes smaller and smaller. In the limit where x goes to infinity, $\frac{1}{x}$ approaches zero: $\lim_{x\to\infty} \frac{1}{x} = 0$. On the other hand, when x takes on smaller and smaller positive values, the expression $\frac{1}{x}$ becomes infinite: $\lim_{x\to 0+} \frac{1}{x} = \infty$. When x approaches 0 from the left, we have $\lim_{x\to 0-} \frac{1}{x} = -\infty$. If these calculations are not clear to you, study the graph of $f(x) = \frac{1}{x}$.

Here are some other examples of limits:

```
>>> limit(sin(x)/x, x, 0)
1
>>> limit(sin(x)**2/x, x, 0)
0
>>> limit(exp(x)/x**100,x,oo) # which is bigger e^x or x^100 ?
oo                            # exp f >> all poly f for big x
```

Derivatives

The derivative function, denoted $f'(x)$, $\frac{d}{dx}f(x)$, $\frac{df}{dx}$, or $\frac{dy}{dx}$, describes the *rate of change* of the function $f(x)$. The SymPy function `diff` computes the derivative of any expression:

```
>>> diff(x**3, x)
3*x**2
```

The differentiation operation knows the product rule $[f(x)g(x)]' = f'(x)g(x) + f(x)g'(x)$, the chain rule $f(g(x))' = f'(g(x))g'(x)$, and the quotient rule $\left[\frac{f(x)}{g(x)}\right]' = \frac{f'(x)g(x)-f(x)g'(x)}{g(x)^2}$:

```
>>> diff( x**2*sin(x), x )
2*x*sin(x) + x**2*cos(x)
>>> diff( sin(x**2), x )
cos(x**2)*2*x
>>> diff( x**2/sin(x), x )
(2*x*sin(x) - x**2*cos(x))/sin(x)**2
```

The second derivative of a function `f` is `diff(f,x,2)`:

```
>>> diff(x**3, x, 2)        # same as diff(diff(x**3, x), x)
6*x
```

Tangent lines

The *tangent line* to the function $f(x)$ at $x = x_0$ is the line that passes through the point $(x_0, f(x_0))$ and has the same slope as the function at that point. The tangent line to the function $f(x)$ at the point $x = x_0$ is described by the equation

$$T_1(x) = f(x_0) + f'(x_0)(x - x_0).$$

What is the equation of the tangent line to $f(x) = \frac{1}{2}x^2$ at $x_0 = 1$?

```
>>> f = S('1/2')*x**2
>>> f
x**2/2
>>> df = diff(f, x)
>>> df
x
>>> T_1 = f.subs({x:1}) + df.subs({x:1})*(x - 1)
>>> T_1
x - 1/2               #  y = x - 1/2
```

The tangent line $T_1(x)$ has the same value and slope as the function $f(x)$ at $x = 1$:

```
>>> T_1.subs({x:1}) == f.subs({x:1})
True
>>> diff(T_1, x).subs({x:1}) == diff(f, x).subs({x:1})
True
```

See Figure 5.10 on page 335.

Optimization

Recall the *second derivative test* for finding the maxima and minima of a function, which we learned on page 351.

Let's find the critical points of the function $f(x) = x^3 - 2x^2 + x$ and use the information from its second derivative to find the maximum of the function on the interval $x \in [0, 1]$.

```
>>> x = Symbol('x')
>>> f = x**3-2*x**2+x
>>> diff(f, x)
3*x**2 - 4*x + 1
>>> sols = solve( diff(f,x),  x)
>>> sols
```

```
[1/3, 1]
>>> diff(diff(f,x), x).subs( {x:sols[0]} )
-2
>>> diff(diff(f,x), x).subs( {x:sols[1]} )
2
```

It will help to look at the graph of this function. The point $x = \frac{1}{3}$ is a local maximum because it is a critical point of $f(x)$ where the curvature is negative, meaning $f(x)$ looks like the peak of a mountain at $x = \frac{1}{3}$. The maximum value of $f(x)$ on the interval $x \in [0,1]$ is $f\left(\frac{1}{3}\right) = \frac{4}{27}$. The point $x = 1$ is a local minimum because it is a critical point with positive curvature, meaning $f(x)$ looks like the bottom of a valley at $x = 1$.

Integrals

In SymPy we use `integrate(f, x)` to obtain the integral function $F(x)$ of any function $f(x)$: $F(x) = \int_0^x f(u)\,du$.

```
>>> integrate(x**3, x)
x**4/4
>>> integrate(sin(x), x)
-cos(x)
>>> integrate(ln(x), x)
x*log(x) - x
```

This is known as an *indefinite integral* since the limits of integration are not defined.

In contrast, a *definite integral* computes the area under $f(x)$ between $x = a$ and $x = b$. Use `integrate(f, (x,a,b))` to compute the definite integrals of the form $A(a,b) = \int_a^b f(x)\,dx$:

```
>>> integrate(x**3, (x,0,1))
1/4                    # the area under x^3 from x=0 to x=1
```

We can obtain the same area by first calculating the indefinite integral $F(c) = \int_0^c f(x)\,dx$, then using $A(a,b) = F(x)\big|_a^b = F(b) - F(a)$:

```
>>> F = integrate(x**3, x)
>>> F.subs({x:1}) - F.subs({x:0})
1/4
```

Integrals correspond to *signed* area calculations:

```
>>> integrate(sin(x), (x,0,pi))
2
>>> integrate(sin(x), (x,pi,2*pi))
-2
>>> integrate(sin(x), (x,0,2*pi))
0
```

During the first half of its 2π-cycle, the graph of $\sin(x)$ is above the x-axis, so it has a positive contribution to the area under the curve. During the second half of its cycle (from $x = \pi$ to $x = 2\pi$), $\sin(x)$ is below the x-axis, so it contributes negative area. Draw a graph of $\sin(x)$ to see what is going on.

Fundamental theorem of calculus

The integral is the "inverse operation" of the derivative. If you perform the integral operation followed by the derivative operation on some function, you'll obtain the same function:

$$\left(\frac{d}{dx} \circ \int dx \right) f(x) = \frac{d}{dx} \int_c^x f(u)\, du = f(x).$$

```
>>> f = x**2
>>> F = integrate(f, x)
>>> F
x**3/3              # + C
>>> diff(F, x)
x**2
```

Alternately, if you compute the derivative of a function followed by the integral, you will obtain the original function $f(x)$ (up to a constant):

$$\left(\int dx \circ \frac{d}{dx} \right) f(x) = \int_c^x f'(u)\, du = f(x) + C.$$

```
>>> f = x**2
>>> df = diff(f, x)
>>> df
2*x
>>> integrate(df, x)
x**2    # + C
```

The fundamental theorem of calculus is important because it tells us how to solve differential equations. If we have to solve for $f(x)$ in the differential equation $\frac{d}{dx} f(x) = g(x)$, we can take the integral on both sides of the equation to obtain the answer $f(x) = \int g(x)\, dx + C$.

Sequences

Sequences are functions that take whole numbers as inputs. Instead of continuous inputs $x \in \mathbb{R}$, sequences take natural numbers $n \in \mathbb{N}$ as inputs. We denote sequences as a_n instead of the usual function notation $a(n)$.

We define a sequence by specifying an expression for its nth term:

```
>>> a_n = 1/n
>>> b_n = 1/factorial(n)
```

Substitute the desired value of n to see the value of the nth term:

```
>>> a_n.subs({n:5})
1/5
```

The Python list comprehension syntax `[item for item in list]` can be used to print the sequence values for some range of indices:

```
>>> [ a_n.subs({n:i}) for i in range(0,8) ]
[oo, 1, 1/2, 1/3, 1/4,  1/5,  1/6,  1/7]
>>> [ b_n.subs({n:i}) for i in range(0,8) ]
[1,  1, 1/2, 1/6, 1/24, 1/120, 1/720, 1/5040]
```

Observe that a_n is not properly defined for $n = 0$ since $\frac{1}{0}$ is a division-by-zero error. To be precise, we should say a_n's domain is the positive naturals $a_n : \mathbb{N}^+ \to \mathbb{R}$. Observe how quickly the `factorial` function $n! = 1 \cdot 2 \cdot 3 \cdots (n-1) \cdot n$ grows: $7! = 5040$, $10! = 3628800$, $20! > 10^{18}$.

We're often interested in calculating the limits of sequences as $n \to \infty$. What happens to the terms in the sequence when n becomes large?

```
>>> limit(a_n, n, oo)
0
>>> limit(b_n, n, oo)
0
```

Both $a_n = \frac{1}{n}$ and $b_n = \frac{1}{n!}$ *converge* to 0 as $n \to \infty$.

Many important math quantities are defined as limit expressions. An interesting example to consider is the number π, which is defined as the area of a circle of radius 1. We can approximate the area of the unit circle by drawing a many-sided regular polygon around the circle. Splitting the n-sided regular polygon into identical triangular splices, we can obtain a formula for its area A_n (see solution to **P1.36**). In the limit as $n \to \infty$, the n-sided-polygon approximation to the area of the unit-circle becomes exact:

```
>>> A_n = n*tan(2*pi/(2*n))
>>> limit(A_n, n, oo)
pi
```

Series

Suppose we're given a sequence a_n and we want to compute the sum of all the values in this sequence $\sum_n^\infty a_n$. Series are sums of sequences. Summing the values of a sequence $a_n : \mathbb{N} \to \mathbb{R}$ is analogous to taking the integral of a function $f : \mathbb{R} \to \mathbb{R}$.

To work with series in `SymPy`, use the `summation` function whose syntax is analogous to the `integrate` function:

```
>>> a_n = 1/n
>>> b_n = 1/factorial(n)
>>> summation(a_n, [n, 1, oo])
oo
>>> summation(b_n, [n, 0, oo])
E
```

We say the series $\sum a_n$ *diverges* to infinity (or *is divergent*) while the series $\sum b_n$ converges (or *is convergent*). As we sum together more and more terms of the sequence b_n, the total becomes closer and closer to some finite number. In this case, the infinite sum $\sum_{n=0}^{\infty} \frac{1}{n!}$ converges to the number $e = 2.71828\ldots$.

The summation command is useful because it allows us to compute *infinite* sums, but for most practical applications we don't need to take an infinite number of terms in a series to obtain a good approximation. This is why series are so neat: they represent a great way to obtain approximations.

Using standard Python commands, we can obtain an approximation to e that is accurate to six decimals by summing 10 terms in the series:

```
>>> import math
>>> def b_nf(n):
        return 1.0/math.factorial(n)
>>> sum( [b_nf(n) for n in range(0,10)] )
2.718281 52557319
>>> E.evalf()
2.718281 82845905        # true value
```

Taylor series

Wait, there's more! Not only can we use series to approximate numbers, we can also use them to approximate functions.

A *power series* is a series whose terms contain different powers of the variable x. The nth term in a power series is a function of both the sequence index n and the input variable x.

For example, the power series of the function $\exp(x) = e^x$ is

$$\exp(x) = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \frac{x^4}{4!} + \frac{x^5}{5!} + \cdots = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

This is, IMHO, one of the most important ideas in calculus: you can compute the value of $\exp(5)$ by taking the infinite sum of the terms in the power series with $x = 5$:

```
>>> exp_xn = x**n/factorial(n)
>>> summation( exp_xn.subs({x:5}), [n, 0, oo] ).evalf()
148.413159102577
```

```
>>> exp(5).evalf()
148.413159102577          # the true value
```

Note that SymPy is actually smart enough to recognize that the infinite series you're computing corresponds to the closed-form expression e^5:

```
>>> summation( exp_xn.subs({x:5}), [n, 0, oo])
exp(5)
```

Taking as few as 35 terms in the series is sufficient to obtain an approximation to e that is accurate to 16 decimals:

```
>>> import math                      # redo using only python
>>> def exp_xnf(x,n):
        return x**n/math.factorial(n)
>>> sum( [exp_xnf(5.0,i) for i in range(0,35)] )
148.413159102577
```

The coefficients in the power series of a function (also known as the *Taylor series*) depend on the value of the higher derivatives of the function. The formula for the nth term in the Taylor series of $f(x)$ expanded at $x = c$ is $a_n(x) = \frac{f^{(n)}(c)}{n!}(x - c)^n$, where $f^{(n)}(c)$ is the value of the nth derivative of $f(x)$ evaluated at $x = c$. The term *Maclaurin series* refers to Taylor series expansions at $x = 0$.

The SymPy function `series` is a convenient way to obtain the series of any function. Calling `series(expr,var,at,nmax)` will show you the series expansion of `expr` near `var=at` up to power `nmax`:

```
>>> series( sin(x), x, 0, 8)
x - x**3/6 + x**5/120 - x**7/5040 + O(x**8)
>>> series( cos(x), x, 0, 8)
1 - x**2/2 + x**4/24 - x**6/720 + O(x**8)
>>> series( sinh(x), x, 0, 8)
x + x**3/6 + x**5/120 + x**7/5040 + O(x**8)
>>> series( cosh(x), x, 0, 8)
1 + x**2/2 + x**4/24 + x**6/720 + O(x**8)
```

Some functions are not defined at $x = 0$, so we expand them at a different value of x. For example, the power series of $\ln(x)$ expanded at $x = 1$ is

```
>>> series(ln(x), x, 1, 6)      # Taylor series of ln(x) at x=1
x - x**2/2 + x**3/3 - x**4/4 + x**5/5  + O(x**6)
```

Here, the result SymPy returns is misleading. The Taylor series of $\ln(x)$ expanded at $x = 1$ has terms of the form $(x - 1)^n$:

$$\ln(x) = (x - 1) - \frac{(x - 1)^2}{2} + \frac{(x - 1)^3}{3} - \frac{(x - 1)^4}{4} + \frac{(x - 1)^5}{5} + \cdots .$$

Verify this is the correct formula by substituting $x = 1$. SymPy returns an answer in terms of coordinates *relative* to $x = 1$.

Instead of expanding $\ln(x)$ around $x = 1$, we can obtain an equivalent expression if we expand $\ln(x + 1)$ around $x = 0$:

```
>>> series(ln(x+1), x, 0, 6)    # Maclaurin series of ln(x+1)
x - x**2/2 + x**3/3 - x**4/4 + x**5/5 + O(x**6)
```

Vectors

A vector $\vec{v} \in \mathbb{R}^n$ is an n-tuple of real numbers. For example, consider a vector that has three components:

$$\vec{v} = (v_1, v_2, v_3) \ \in \ \mathbb{R}^3.$$

To specify the vector $\vec{v}$, we specify the values for its three components v_1, v_2, and v_3.

A matrix $A \in \mathbb{R}^{m \times n}$ is a rectangular array of real numbers with m rows and n columns. A vector is a special type of matrix; you can think of the vector $\vec{v} \in \mathbb{R}^n$ as a $1 \times n$ matrix. Because of this equivalence between vectors and matrices, in SymPy we use Matrix objects to represent vectors.

This is how we define vectors and compute their properties:

```
>>> u = Matrix([4,5,6])
>>> u
[4, 5, 6]                       # 3-vector
>>> u[0]                        # 0-based indexing for entries
4
>>> u.norm()                    # length of u
sqrt(77)
>>> uhat = u/u.norm()           # unit vector in same dir as u
>>> uhat
[4/sqrt(77), 5/sqrt(77), 6/sqrt(77)]
>>> uhat.norm()
1
```

Dot product

The dot product of the 3-vectors $\vec{u}$ and $\vec{v}$ can be defined two ways:

$$\vec{u} \cdot \vec{v} = \underbrace{u_x v_x + u_y v_y + u_z v_z}_{\text{algebraic def.}} = \underbrace{\|\vec{u}\|\|\vec{v}\|\cos(\varphi)}_{\text{geometric def.}} \ \in \mathbb{R},$$

where φ is the angle between the vectors $\vec{u}$ and $\vec{v}$. In SymPy,

```
>>> u = Matrix([ 4,5,6])
>>> v = Matrix([-1,1,2])
>>> u.dot(v)
13
```

We can combine the algebraic and geometric formulas for the dot product to obtain the cosine of the angle between the vectors

$$\cos(\varphi) = \frac{\vec{u} \cdot \vec{v}}{\|\vec{u}\|\|\vec{v}\|} = \frac{u_x v_x + u_y v_y + u_z v_z}{\|\vec{u}\|\|\vec{v}\|},$$

and use the `acos` function to find the angle measure:

```
>>> acos(u.dot(v)/(u.norm()*v.norm())).evalf()
0.921263115666387      # in radians  =  52.76 degrees
```

Just by looking at the coordinates of the vectors $\vec{u}$ and $\vec{v}$, it's difficult to determine their relative direction. Thanks to the dot product, however, we know the angle between the vectors is $52.76°$, which means they *kind of* point in the same direction. Vectors that are at an angle $\varphi = 90°$ are called *orthogonal*, meaning at right angles with each other. The dot product between two vectors is negative when the angle between them is $\varphi > 90°$.

The notion of the "angle between vectors" applies more generally to vectors with any number of dimensions. The dot product for n-dimensional vectors is $\vec{u} \cdot \vec{v} = \sum_{i=1}^{n} u_i v_i$. This means we can talk about "the angle between" 1000-dimensional vectors. That's pretty crazy if you think about it—there is no way we could possibly "visualize" 1000-dimensional vectors, yet given two such vectors we can tell if they point mostly in the same direction, in perpendicular directions, or mostly in opposite directions.

The dot product is a commutative operation $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$:

```
>>> u.dot(v) == v.dot(u)
True
```

Cross product

The *cross product*, denoted $\times$, takes two vectors as inputs and produces a vector as output. The cross products of individual basis elements are defined as follows:

$$\hat{\imath} \times \hat{\jmath} = \hat{k}, \qquad \hat{\jmath} \times \hat{k} = \hat{\imath}, \qquad \hat{k} \times \hat{\imath} = \hat{\jmath}.$$

The cross product is defined by the following equation:

$$\vec{u} \times \vec{v} = \left(u_y v_z - u_z v_y, \ u_z v_x - u_x v_z, \ u_x v_y - u_y v_x\right).$$

Here's how to compute the cross product of two vectors:

```
>>> u = Matrix([ 4,5,6])
>>> v = Matrix([-1,1,2])
>>> u.cross(v)
[4, -14, 9]
```

The vector $\vec{u} \times \vec{v}$ is orthogonal to both $\vec{u}$ and $\vec{v}$. The norm of the cross product $\|\vec{u} \times \vec{v}\|$ is proportional to the lengths of the vectors and the sine of the angle between them:

```
(u.cross(v).norm()/(u.norm()*v.norm()))).n()
0.796366206088088     # = sin(0.921..)
```

The cross product is anticommutative, $\vec{u} \times \vec{v} = -\vec{v} \times \vec{u}$:

```
>>> u.cross(v)
[4, -14, 9]
>>> v.cross(u)
[-4, 14,-9]
```

Watch out for this, because it's a new thing. The product of two numbers a and b is commutative: $ab = ba$. The dot product of two vectors $\vec{u}$ and $\vec{v}$ is commutative: $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$. However the cross product is not commutative: $\vec{u} \times \vec{v} \neq \vec{v} \times \vec{u}$, it is anticommutative: $\vec{u} \times \vec{v} = -\vec{v} \times \vec{u}$.

Mechanics

The module called `sympy.physics.mechanics` contains elaborate tools for describing mechanical systems, manipulating reference frames, forces, and torques. These specialized functions are not necessary for a first-year mechanics course. The basic `SymPy` functions like `solve`, and the vector operations you learned in the previous sections are powerful enough for basic Newtonian mechanics.

Dynamics

The net force acting on an object is the sum of all the external forces acting on it $\vec{F}_{net} = \sum \vec{F}$. Since forces are vectors, we need to use vector addition to compute the net force.

Compute $\vec{F}_{net} = \vec{F}_1 + \vec{F}_2$, where $\vec{F}_1 = 4\hat{\imath}$[N] and $\vec{F}_2 = 5\angle 30°$[N]:

```
>>> F_1 = Matrix( [4,0] )
>>> F_2 = Matrix( [5*cos(30*pi/180), 5*sin(30*pi/180)] )
>>> F_net = F_1 + F_2
>>> F_net
[4 + 5*sqrt(3)/2, 5/2]
>>> F_net.evalf()
[8.33012701892219, 2.5]                    # in Newtons
```

To express the answer in length-and-direction notation, use `norm` to find the length of $\vec{F}_{net}$, and use the two-input inverse tangent function `atan2` to compute the direction:

```
>>> F_net.norm().evalf()
8.69718438067042                              # |F_net| in [N]
>>> (atan2(F_net[1],F_net[0])*180/pi).n()
16.7053138060100                              # angle in degrees
```

The net force on the object is $\vec{F}_{net} = 8.697\angle 16.7°$[N].

Kinematics

Let $x(t)$ denote the position of an object, $v(t)$ denote its velocity, and $a(t)$ denote its acceleration. Together $x(t)$, $v(t)$, and $a(t)$ are known as the *equations of motion* of the object.

Starting from the knowledge of $\vec{F}_{net}$, we can compute $a(t) = \frac{\vec{F}_{net}}{m}$, then obtain $v(t)$ by integrating $a(t)$, and finally obtain $x(t)$ by integrating $v(t)$:

$$\underbrace{\frac{\vec{F}_{net}}{m} = a(t)}_{\text{Newton's 2}^{\text{nd}}\text{ law}} \underbrace{\xrightarrow{v_i + \int dt} v(t) \xrightarrow{x_i + \int dt} x(t).}_{\text{kinematics}}$$

Uniform acceleration motion (UAM)

Let's analyze the case where the net force on the object is constant. A constant force causes a constant acceleration $a = \frac{F}{m}$ = constant. If the acceleration function is constant over time $a(t) = a$. We find $v(t)$ and $x(t)$ as follows:

```
>>> t, a, v_i, x_i = symbols('t a v_i x_i')
>>> v = v_i + integrate(a, (t,0,t))
>>> v
a*t + v_i
>>> x = x_i + integrate(v, (t,0,t))
>>> x
a*t**2/2 + v_i*t + x_i
```

You may remember these equations from Section 2.4 (page 170). They are the *uniform accelerated motion* (UAM) equations:

$$\begin{aligned} a(t) &= a, \\ v(t) &= v_i + at, \\ x(t) &= x_i + v_i t + \frac{1}{2}at^2. \end{aligned}$$

In high school, you probably had to memorize these equations. Now you know how to derive them yourself starting from first principles.

For the sake of completeness, we'll now derive the fourth UAM equation, which relates the object's final velocity to the initial velocity, the displacement, and the acceleration, without reference to time:

```
>>> expand(v*v)
a**2*t**2 + 2*a*t*v_i + v_i**2
>>> simplify(expand(v*v) - 2*a*x)
-2*a*x_i + v_i**2
```

The above calculation shows $v_f^2 - 2ax_f = -2ax_i + v_i^2$. After moving the term $2ax_f$ to the other side of the equation, we obtain

$$(v(t))^2 \;=\; v_f^2 = v_i^2 + 2a\Delta x \;=\; v_i^2 + 2a(x_f - x_i).$$

The fourth equation is important for practical purposes because it allows us to solve physics problems without using the time variable.

Example

Find the position function of an object at time $t = 3$[s], if it starts from $x_i = 20$[m] with $v_i = 10$[m/s] and undergoes a constant acceleration of $a = 5$[m/s^2]. What is the object's velocity at $t = 3$[s]?

```
>>> x_i = 20    # initial position
>>> v_i = 10    # initial velocity
>>> a = 5       # acceleration (constant during motion)
>>> v = v_i + integrate(a, (t,0,t))
>>> x = x_i + integrate(v, (t,0,t))
>>> x
5*t**2/2 + 10*t + 20
>>> x.subs({t:3}).n()           # x(3) in [m]
72.5
>>> diff(x,t).subs({t:3}).n()   # v(3) in [m/s]
25                              # = sqrt( v_i**2 + 2*a*52.5 )
```

If you think about it, physics knowledge combined with computer skills is like a superpower!

General equations of motion

The procedure $a(t) \xrightarrow{v_i + \int dt} v(t) \xrightarrow{x_i + \int dt} x(t)$ can be used to obtain the position function $x(t)$ even when the acceleration is not constant. Suppose the acceleration of an object is $a(t) = \sqrt{kt}$; what is its $x(t)$?

```
>>> t, v_i, x_i, k = symbols('t v_i x_i k')
>>> a = sqrt(k*t)
>>> v = v_i + integrate(a, (t,0,t))
>>> x = x_i + integrate(v, (t,0,t))
>>> x
x_i + v_i*t + (4/15)*(k*t)**(5/2)/k**2
```

Potential energy

For each force $\vec{F}(x)$ there is a corresponding potential energy $U_F(x)$. The change in potential energy associated with the force $\vec{F}(x)$ and displacement $\vec{d}$ is defined as the negative of the work done by the force during the displacement: $U_F(x) = -W = -\int_{\vec{d}} \vec{F}(x) \cdot d\vec{x}$.

The potential energies associated with gravity $\vec{F}_g = -mg\hat{j}$ and the force of a spring $\vec{F}_s = -k\vec{x}$ are calculated as follows:

```
>>> x, y = symbols('x y')
>>> m, g, k, h = symbols('m g k h')

>>> F_g = -m*g                  # Force of gravity on mass m
>>> U_g = - integrate(F_g, (y,0,h))
>>> U_g
m*g*h                           # Grav. potential energy

>>> F_s = -k*x                  # Spring force for displacement x
>>> U_s = - integrate(F_s, (x,0,x))
>>> U_s
k*x**2/2                        # Spring potential energy
```

Note the negative sign in the formula defining the potential energy. This negative is cancelled by the negative sign of the dot product $\vec{F} \cdot d\vec{x}$: when the force acts in the direction opposite to the displacement, the work done by the force is negative.

Simple harmonic motion

```
from sympy import Function, dsolve
```

The force exerted by a spring is given by the formula $F = -kx$. If the only force acting on a mass m is the force of a spring, we can use Newton's second law to obtain the following equation:

$$F = ma \quad \Rightarrow \quad -kx = ma \quad \Rightarrow \quad -kx(t) = m\frac{d^2}{dt^2}\Big[x(t)\Big].$$

The motion of a mass-spring system is described by the *differential equation* $\frac{d^2}{dt^2}x(t) + \omega^2 x(t) = 0$, where the constant $\omega = \sqrt{\frac{k}{m}}$ is called the angular frequency. We can find the position function $x(t)$ using the dsolve method:

```
>>> t = Symbol('t')                 # time            t
>>> x = Function('x')               # position function x(t)
>>> w = Symbol('w', positive=True)  # angular frequency w
>>> sol = dsolve( diff(x(t),t,t) + w**2*x(t), x(t) )
>>> sol
```

```
x(t) == C1*sin(w*t) + C2*cos(w*t)
>>> x = sol.rhs
>>> x
C1*sin(w*t) + C2*cos(w*t)
```

Note the solution $x(t) = C_1 \sin(\omega t) + C_2 \cos(\omega t)$ is equivalent to $x(t) = A \cos(\omega t + \phi)$, which is more commonly used to describe simple harmonic motion. We can use the expand function with the argument trig=True to convince ourselves of this equivalence:

```
>>> A, phi = symbols("A phi")
>>> expand(A*cos(w*t - phi), trig=True)
A*sin(phi)*sin(w*t) + A*cos(phi)*cos(w*t)
```

If we define $C_1 = A \sin(\phi)$ and $C_2 = A \cos(\phi)$, we obtain the form $x(t) = C_1 \sin(\omega t) + C_2 \cos(\omega t)$ that SymPy found.

Conservation of energy

We can verify that the total energy of the mass-spring system is conserved by showing $E_T(t) = U_s(t) + K(t) = \text{constant}$:

```
>>> x = sol.rhs.subs({"C1":0,"C2":A})
>>> x
A*cos(t*w)
>>> v = diff(x, t)
-A*w*sin(t*w)
>>> E_T = (0.5*k*x**2 + 0.5*m*v**2).simplify()
>>> E_T
0.5*A**2*(k*cos(w*t)**2 + m*w**2*sin(w*t)**2)
>>> E_T.subs({k:m*w**2}).simplify()
0.5*m*(w*A)**2                          # = K_max
>>> E_T.subs({w:sqrt(k/m)}).simplify()
0.5*k*A**2                              # = U_max
```

Conclusion

I'll conclude with some words of warning about technological dependence. Computer technology is very powerful and everywhere around us, but we must not forget that computers are merely calculators that depend on human commands to direct them. This means it's important for you to learn how to do math "by hand" first, in order to know how to instruct computers to do the math for you.

I don't want you to use the tricks you learned in this tutorial to avoid learning how to do math and blindly rely on SymPy to do math calculations for you. That's not a good idea! I'll be very disappointed if you use SymPy to skip the "intellectual suffering" necessary to learn the new math concepts like numbers, equations, functions, etc. That's what math is all about—understand math concepts

and the relationships between concepts. The part that is about rote memorized of math calculations procedures that you should have learned at school is not important at all. The tedious and repetitive math calculations is precisely what can be "outsourced" to `SymPy`.

To solve problems in math (or physics, chemistry, biology, etc.) the most important things are to: A) define the variables relevant for the problem, B) draw a diagram, and C) clearly set up the problem's equations in terms of the variables you defined. With these steps in place, half the work of solving the problem is already done! Computers can't help with these important, initial modelling and problem-specific tasks—only humans are good at this stuff. Once you set up the problem (A, B, C), `SymPy` can help you breeze through any subsequent calculations that might be necessary to obtain the final answer.

Most of the big math and science discoveries were made using pen and paper, which shows that scribbling on paper is a useful as a tool for thinking. With what you learned about `SymPy`, you now have access to the combination of pencil and paper for thinking and `SymPy` for calculating. It's a very powerful combination! What is a real-world problem you'd like to solve? Try modelling the problem using math equations and see what happens. Go out there and do some science!

Links

[Installation instructions for `jupyter notebook`]
`https://jupyter.readthedocs.io/en/latest/install.html`

[The official `SymPy` tutorial]
`http://docs.sympy.org/latest/tutorial/intro.html`

[A list of `SymPy` gotchas]
`http://docs.sympy.org/dev/gotchas.html`

Appendix E

Formulas

Calculus formulas

$\dfrac{dy}{dx}$ $\longleftarrow$	y	$\longrightarrow$ $\displaystyle\int y\,dx$
Algebraic		
1	x	$\frac{1}{2}x^2 + C$
0	a	$ax + C$
nx^{n-1}	x^n	$\dfrac{1}{n+1}x^{n+1} + C$
$-x^{-2}$	x^{-1}	$\ln x + C$
$\dfrac{du}{dx} \pm \dfrac{dv}{dx} \pm \dfrac{dw}{dx}$	$u \pm v \pm w$	$\int u\,dx \pm \int v\,dx \pm \int w\,dx$
$u\dfrac{dv}{dx} + v\dfrac{du}{dx}$	uv	No general form known
$\dfrac{v\dfrac{du}{dx} - u\dfrac{dv}{dx}}{v^2}$	$\dfrac{u}{v}$	No general form known
	u	$ux - \int x\,du + C$

$\dfrac{dy}{dx}$	$\longleftarrow \quad y \quad \longrightarrow$	$\displaystyle\int y\,dx$
Exponential and Logarithmic		
e^x	e^x	$e^x + C$
x^{-1}	$\ln x$	$x(\ln x - 1) + C$
$\frac{1}{\ln 10}\,x^{-1}$	$\log_{10} x$	$\frac{1}{\ln 10}\,x(\ln x - 1) + C$
$a^x \ln a$	a^x	$\dfrac{a^x}{\ln a} + C$
Trigonometric		
$\cos x$	$\sin x$	$-\cos x + C$
$-\sin x$	$\cos x$	$\sin x + C$
$\sec^2 x$	$\tan x$	$-\ln|\cos x| + C$
Inverse trigonometric		
$\dfrac{1}{\sqrt{(1 - x^2)}}$	$\sin^{-1}(x)$	$x\sin^{-1}(x) + \sqrt{1 - x^2} + C$
$-\dfrac{1}{\sqrt{(1 - x^2)}}$	$\cos^{-1}(x)$	$x\cos^{-1}(x) - \sqrt{1 - x^2} + C$
$\dfrac{1}{1 + x^2}$	$\tan^{-1}(x)$	$x\tan^{-1}(x) - \frac{1}{2}\ln(1 + x^2) + C$
Hyperbolic		
$\cosh x$	$\sinh x$	$\cosh x + C$
$\sinh x$	$\cosh x$	$\sinh x + C$
$\operatorname{sech}^2 x$	$\tanh x$	$\ln(\cosh x) + C$
Inverse hyperbolic		
$-\dfrac{x}{(a^2 + x^2)^{\frac{3}{2}}}$	$\dfrac{1}{\sqrt{a^2 + x^2}}$	$\sinh^{-1}(\frac{x}{a}) + C = \ln(x + \sqrt{a^2 + x^2}) + C$

$\dfrac{dy}{dx}$	$\longleftarrow \quad y \quad \longrightarrow$	$\displaystyle\int y\,dx$
Miscellaneous		
$-\dfrac{1}{(x-a)^2}$	$\dfrac{1}{x-a}$	$\ln|x-a| + C, \quad x \neq a$
$\mp\dfrac{b}{(a\pm bx)^2}$	$\dfrac{1}{a\pm bx}$	$\pm\dfrac{1}{b}\ln|a\pm bx| + C, \quad a\pm bx \neq 0$
$-\dfrac{3a^2 x}{(a^2+x^2)^{\frac{5}{2}}}$	$\dfrac{a^2}{(a^2+x^2)^{\frac{3}{2}}}$	$\dfrac{x}{\sqrt{a^2+x^2}} + C$
$a\cos ax$	$\sin ax$	$-\dfrac{1}{a}\cos ax + C$
$-a\sin ax$	$\cos ax$	$\dfrac{1}{a}\sin ax + C$
$a\sec^2 ax$	$\tan ax$	$-\dfrac{1}{a}\ln|\cos ax| + C$
$\sin 2x$	$\sin^2 x$	$\dfrac{x}{2} - \dfrac{\sin 2x}{4} + C$
$-\sin 2x$	$\cos^2 x$	$\dfrac{x}{2} + \dfrac{\sin 2x}{4} + C$
$n\sin^{n-1}x\cos x$	$\sin^n x$	$-\dfrac{\cos x}{n}\sin^{n-1}x + \dfrac{n-1}{n}\displaystyle\int \sin^{n-2}x\,dx + C$
$-\dfrac{\cos x}{\sin^2 x}$	$\dfrac{1}{\sin x}$	$\ln\left|\tan\dfrac{x}{2}\right| + C$
$-\dfrac{\sin 2x}{\sin^4 x}$	$\dfrac{1}{\sin^2 x}$	$-\cotan x + C$
$\dfrac{\sin^2 x - \cos^2 x}{\sin^2 x \cos^2 x}$	$\dfrac{1}{\sin x \cos x}$	$\ln|\tan x| + C$
$\begin{smallmatrix} n\sin mx\cos nx \;+ \\ m\sin nx\cos mx \end{smallmatrix}$	$\sin mx \sin nx$	$\tfrac{1}{2}\cos(m-n)x - \tfrac{1}{2}\cos(m+n)x + C$
$2a\sin 2ax$	$\sin^2 ax$	$\dfrac{x}{2} - \dfrac{\sin 2ax}{4a} + C$
$-2a\sin 2ax$	$\cos^2 ax$	$\dfrac{x}{2} + \dfrac{\sin 2ax}{4a} + C$

Mechanics formulas

Forces:

$$W = F_g = \frac{GMm}{r^2} = gm, \quad F_s = -kx, \quad F_{fs} \leq \mu_s N, \quad F_{fk} = \mu_k N$$

Newton's three laws:

$$\text{if no } \vec{F}_{\text{ext}}, \text{ then } \vec{v}_i = \vec{v}_f \tag{1}$$

$$\vec{F}_{\text{net}} = m\vec{a} \tag{2}$$

$$\text{if } \vec{F}_{12}, \text{ then } \exists \vec{F}_{21} = -\vec{F}_{12} \tag{3}$$

Uniform acceleration motion (UAM):

$$a(t) = a \tag{4}$$

$$v(t) = at + v_i \tag{5}$$

$$x(t) = \tfrac{1}{2}at^2 + v_i t + x_i \tag{6}$$

$$v_f^2 = v_i^2 + 2a\Delta x \tag{7}$$

Momentum:

$$\vec{p} = m\vec{v} \tag{8}$$

Energy and work:

$$K = \tfrac{1}{2}mv^2, \quad U_g = mgh, \quad U_s = \tfrac{1}{2}kx^2, \quad K_r = \tfrac{1}{2}I\omega^2, \quad W = \vec{F}\cdot\vec{d} \tag{9}$$

Conservation laws:

$$\sum \vec{p}_{\text{in}} = \sum \vec{p}_{\text{out}} \tag{10}$$

$$L_{\text{in}} = L_{\text{out}} \tag{11}$$

$$\sum E_{\text{in}} + W_{\text{in}} = \sum E_{\text{out}} + W_{\text{out}} \tag{12}$$

Circular motion (radial acceleration and radial force):

$$a_r = \frac{v_t^2}{R}, \qquad \vec{F}_r = ma_r \hat{r} \tag{13}$$

Angular motion:

$$F = ma \quad \Rightarrow \quad \mathcal{T} = I\alpha \tag{14}$$

$$a(t), v(t), x(t) \quad \Rightarrow \quad \alpha(t), \omega(t), \theta(t) \tag{15}$$

$$\vec{p} = m\vec{v} \quad \Rightarrow \quad L = I\omega \tag{16}$$

$$K = \tfrac{1}{2}mv^2 \quad \Rightarrow \quad K_r = \tfrac{1}{2}I\omega^2 \tag{17}$$

SHM with $\omega = \sqrt{\frac{k}{m}}$ (mass-spring system) or $\omega = \sqrt{\frac{g}{\ell}}$ (pendulum):

$$x(t) = A\cos(\omega t + \phi) \tag{18}$$

$$v(t) = -A\omega \sin(\omega t + \phi) \tag{19}$$

$$a(t) = -A\omega^2 \cos(\omega t + \phi) \tag{20}$$

Index

The page numbers shown in **bold** point to the concept definitions.

Often calculus and mechanics are taught as separate subjects. It shouldn't be like that. Learning calculus without mechanics is incredibly boring. Learning mechanics without calculus is missing the point. This textbook integrates both subjects and highlights the profound connections between them.

This is the deal. Give me 350 pages of your attention, and I'll teach you everything you need to know about functions, limits, derivatives, integrals, vectors, forces, and accelerations. The book in your hands is the only math book you'll need for the first semester of undergraduate studies in science.

With concise, jargon-free lessons on topics in math and physics, each section covers one concept at the level required for a first-year university course. **Anyone can pick up this book and become proficient in calculus and mechanics, regardless of their mathematical background.**

"I like the conversational tone of your writing. It's almost like I'm learning math from a friend." —Kevin Del Castillo, student

"It's literally two heavy textbooks packed into a very light and small book. My only problem has been when I am working through it in the library and start laughing out loud—you have great examples." —Miriam Aczel, student

The author has more than 14 years of tutoring experience, a B.Eng. in electrical engineering, a M.Sc. in physics, and a Ph.D. in computer science.

Contents:

- HIGH SCHOOL MATH
- VECTORS
- MECHANICS
- DERIVATIVES
- INTEGRALS

The book **No Bullshit Guide to Math and Physics** by Ivan Savov (Minireference Publishing, v5.4 2020, ISBN 0992001005) is available in both print and digital formats.

- Softcover print from lulu.com: 📘 bit.ly/noBSmathphys-sc
- Hardcover print from lulu.com: 📘 bit.ly/noBSmathphys-hc
- Softcover print from amazon: a amazon.com/dp/0992001005
- Digital download from gumroad: 🇬 gum.co/noBSmathphys

For more info, visit the book's website minireference.com.