Compact Calculus - Calculus for the practitioners

Puripat Thumbanthu, Kunakorn Chaiyara

Started: 07/28/2022

Revision: October 7, 2024

Prerequisites: *set theory, algebra, geometry, basic trigonometry*

Contents

Preface

Purpose of the textbook

I probably made a lot of mathematicians angry by writing this book. But it misses their points. This textbook is designed for the practitioners of calculus: people that are interested in fields of sciences including biology, chemistry, and physics. Therefore, expect what you aren't expected, and don't expect what you traditionally expect.

The contents of this textbook are entirely self-contained and organized into four parts (volumes)[1]: dynamical systems and ordinary differential equations, signal analysis, real and complex analysis, and vector calculus with partial differential equations. However, the topics deviates greatly from the traditional way calculus is taught. I took a great care of how the material is arranged in better needs practitioners, which means it may deviate from a standard calculus curriculum. To maintain focus within the core chapters, the auxilaries topics typically covered in traditional courses have been moved to the appendices.

Even though the book is very self-contained, it still doesn't serve as a standalone book. Sure, the story that I'm about to tell is quite fruitful,

[1]The exact number of parts may change as more material is added during the writing process.

and you can enjoy it from the beginning to end. But, it lacks one thing: problem solving experiences. As of the revision today (October 7, 2024), there are not that many exercises in the chapters. So, you might have to find some online. I've scattered plenty of them throughout the chapters.

Layout of the book

Layout of Part I

The first part of the book concerns differential equations: a powerful tool that can be used to describe how a system changes over time. The contents are laid out as follows:

- Chapters 1 and 2 develops the basic building blocks of calculus: derivatives and integrals, from the ground up
- Chapter 3 takes the reader through derivatives and integrals of basic functions, e.g., exponentials, and logarithms.
- Chapter 4 takes a detour from the main objective a bit, and looks at how calculus could be applied to solve various geometrical problems.
- Chapter 5 studies the interaction between calculus and various trigonometric functions. It also applies calculus to describe some physical systems, specifically oscillatory ones. From this chapter, most of the foundations of calculus is already laid down, ready to be used.
- Finally, chapter 6 uses all the knowledge developed in earlier chapters to explore many systems that exist in nature, including biological, chemical, and physical systems.

Acknowledgement

Part I

The foundation of calculus: differential equations.

CHAPTER 1

Writing down nature: derivatives

Prerequisites: *graphs, functions, kinematic variables*

1.1 Invitation to calculus from a practitioner

Calculus is the language of change, whether it's changes in weather patterns, movement of an object, evolution of biological systems, or even changes in your relationship! Although the last one might be diabolical to think about for some of you, calculus provides a mathematical framework to describe and predict these systems. Therefore, I've structured this volume around the concept of dynamical systems: systems that evolve through time, whether they are biological, chemical, or physical.

As said in the preface, I have thrown most unrelated contents into the appendices. But as much as I'd love to do, some purely mathematical contents are still essential for understanding and solving dynamical systems. While I've kept those in had to keep those in, I've tried to make them as brief as possible.

The problem of dynamical systems are separated into two parts:

1. Describing the system, and
2. Solving for the system's future.

Calculus is also separated into two parts: differential and integral. These parts nicely lines up with our problem of describing a dynamical system. Differential calculus uses differential equations to describe a system; those equations are then solved using integral calculus to predict the system's future. Therefore, this chapter is devoted to the first part of calculus: differential calculus.

At the end, you should develop the intuition that

1. Derivative measures the rate of change of a function w.r.t. a variable.
2. Derivatives can be thought of the slope of a graph.
3. The universe is described in the language of differential equations

1.1.1 The notation of differential calculus

If a is a variable, then $\mathrm{d}a$ is a very small quantity of a a.k.a. an **infinitesimal**. E.g., if $\mathbf{x}$ is displacement, then $\mathrm{d}\mathbf{x}$ is a very small displacement. If t is the time, then $\mathrm{d}t$, a very short time.

We can also find ratio between infinitesimal. E.g., the ratio between some small distance and some short time, $\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}$, we get the speed $\mathbf{v}$.

1.2 Speed and instantaneous rate of change

Our first dynamical system that I'd like to discuss is the car. Car moves, and the **speedometer** displays its speed: a change of position through time. How does the speedometer actually work?

If we're traveling at a speed $5\,\mathrm{m\,s^{-1}}$, when *exactly* are we traveling at $5\,\mathrm{m\,s^{-1}}$? You could say $\mathbf{v} = 5\,\mathrm{m}$ *at the moment of measurement*. That's like saying, "Oh, I can find the speed of the car by just taking a picture of it." But that's illegal! To calculate speed, we have to compare two points in space through time. Or, the rate of change of distance through time:

$$\mathbf{v} = \frac{\mathbf{s}_2 - \mathbf{s}_1}{t_2 - t_1}. \tag{1.1}$$

While it may seem like cameras grab snapshots in an instant, they actually need time to take in light to construct an image, they need some *exposure time* Δt.

To get a "not blurred" image of a moving object, we reduce the exposure time. If the exposure time is too long, the object will be smeared out. This effect is known as motion blur, which is normally undesired. But motion blur actually helps out a lot with measuring velocity. In which the velocity is just

$$\mathbf{v} = \frac{\text{Distance between the smear and the main object}}{\text{Exposure time}}. \tag{1.2}$$

As illustrated in fig. 1.1, the motion blur clearly shows us the change in position of the car over the exposure time.

But what will happen with shorter exposure time? Does the motion blur disappear? No! The blur is still there, but it's just smaller. Typically, $12\,\mathrm{ms}$ exposure time is short enough to create a "focused image". But it's just an illusion that came from the limitation of the screen's ability to

Fig. 1.1 | Calculating the velocity of a car from motion blur.

reproduce such little blur. If our camera and screen is good enough, we can *always* calculate the velocity of the car from the blur. The smaller the exposure time, the more detailed the image is, and the closer you'll get to the exact $\mathbf{v}$ at that moment in time. Finally, if we let the exposure time become infinitesimally short, we can say the $\mathbf{v}$ we got is *the velocity at that exact point*. Or as we call it, the **instantaneous velocity**. By using the said calculus notation in section 1.1.1, we can just write this as

$$\mathbf{v} = \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}, \tag{1.3}$$

and here it is ladies and gentlemen, the **derivative**: the measure of rate of change.

From here on out, I shall use *derivative* and *rate of change* interchangeably. So, every time you see *derivative*, think *rate of change*.

1.3 An attempt to define derivatives

Mathematically, a **derivative** is a measure of a function's rate of change with respect to a variable. In the previous example, $\mathbf{x}$ is a function that's dependent on time, and its derivative w.r.t. time, $\mathbf{v}$, is measuring the rate of change of function $\mathbf{x}(t)$ through time $\mathrm{d}t$.

To find an explicit expression for derivative, let's say we have two points in spacetime $(\mathbf{x}_1, t_1)$ and $(\mathbf{x}_2, t_2)$. The change in $\mathbf{x}$ is $\mathbf{x}_2 - \mathbf{x}_1$, and the change in time is $t_2 - t_1$. The derivative of position w.r.t. time is then just

$$\mathbf{v} = \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \frac{\mathbf{x}_2 - \mathbf{x}_1}{t_2 - t_1}. \tag{1.4}$$

Let $t_2 = t + h$. Then, $\mathbf{x}_2 = \mathbf{x}(t+h)$ and $\mathbf{x}_1 = \mathbf{x}(t)$. The time difference used to calculate $\mathbf{v}$ must be miniscule: infinitesimally small. I have to introduce the notion of limits, which is just a fancier way of saying "very close to, but not"

$$\mathbf{v} = \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \lim_{h \to 0} \frac{\mathbf{x}(t+h) - \mathbf{x}(t)}{(t+h) - t} \tag{1.5}$$

$$= \lim_{h \to 0} \frac{\mathbf{x}(t+h) - \mathbf{x}(t)}{h}. \tag{1.6}$$

The equation above is what we generally refer to as the *definition of derivative*. Then, we just extend this relation to any function $f(x)$, which requires just a substitution of variables. And then we get:

Definition 1: Naive definition of derivative

A derivative of a function $f(x)$ w.r.t. a variable x is the rate of change of $f(x)$ w.r.t. x, and it is written as

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} \quad \text{or,} \quad \frac{\mathrm{d}}{\mathrm{d}x} f(x), \tag{1.7}$$

where

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}. \tag{1.8}$$

Equation (1.8) directly reads:

The derivative of $f(x)$ with respect to x is $\dfrac{f(x+h) - f(x)}{h}$ where h is very close to 0, but h is strictly not 0.

On notation: So far, we've been using the Leibniz's notation for derivatives, and it has the property that derivatives behave exactly like fractions, and you can cancel terms.

$$\frac{\mathrm{d}a}{\mathrm{d}b} \times \frac{\mathrm{d}b}{\mathrm{d}c} = \frac{\mathrm{d}a}{\mathrm{d}c}. \tag{1.9}$$

But, this isn't the only accepted notation. Multiple great mathematicians have come up with their own, and some are better than others in certain cases. I'd introduce other notations later on if necessary.

1.4 The geometrical interpretation of the derivative

One way to interpret derivatives is by using slope. Notice that eq. (1.4) looks a lot like the slope equation

$$m = \frac{y_2 - y_1}{x_2 - x_1}. \tag{1.10}$$

We've also seen that eq. (1.4) is analogous to the definition of derivative eq. (1.8). So is the derivative just the slope of a line? If it is, then what is the y-axis and the x-axis of a graph?

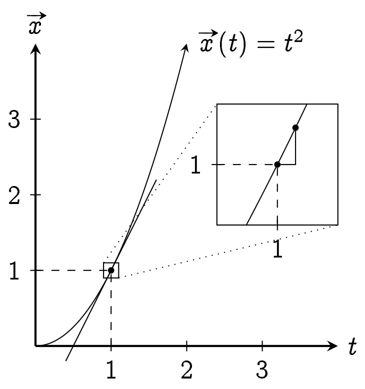


FIG. 1.2 | POSITION VS. TIME GRAPH WHERE $\mathbf{x}(t) = t^2$.

We can compare eq. (1.10) to eq. (1.4): the y-axis should be the position $\mathbf{x}$, and the x-axis, the time t. If we draw that out, we'll get the x-t graph, which can encode the exact trajectory of an object. An example of which is shown in fig. 1.2.

But what does this have to do with derivatives? Equation (1.10) only works for straight line! Well, here's the beauty of it. If you zoom into any points on a curve, eventually, it will look like a line. And thus, *the derivative zooms into the curve at some point, and chooses two very close points on the curve and calculate its slope. In which, that slope represents the rate of change of the function at that point.*

1.5 Evaluation of derivatives: method of increments

The derivative's definition can be used to directly evaluate derivatives. This is called the **method of increment**. E.g., in fig. 1.2, let's evaluate the velocity at time $t = 1\,\mathrm{s}$ where the position function, $\mathbf{x}(t) = t^2$. $t = 1\,\mathrm{s}$. We start from eq. (1.8):

$$\begin{aligned}
\mathbf{v}(1) = \frac{\mathbf{x_2} - \mathbf{x_1}}{t_2 - t_1} &= \lim_{h \to 0} \frac{\mathbf{x}(1+h) - \mathbf{x}(1)}{(1+h) - 1} \\
&= \lim_{h \to 0} \frac{(1+h)^2 - 1}{h} \\
&= \lim_{h \to 0} \frac{2h + h^2}{h} \\
&= \lim_{h \to 0} 2 + h .
\end{aligned}$$

Since h is very close to zero, we approximate $2 + h$ as 2. Imagine comparing 2 to 10^{-10}. The 10^{-10} wouldn't make a noticeable difference, and we can ignore it. Therefore, $\mathbf{v}(t = 1\,\mathrm{s}) = 2\,\mathrm{m\,s^{-1}}$.

Now, try evaluating $\mathbf{v}(t = 3\,\mathrm{s})$ for $\mathbf{x}(t) = t^3$. You should get $81\,\mathrm{m\,s^{-1}}$. As a hint, you can also ignore h^2 because if $h < 1$, then $h^2 < h$.

1.6 Higher order derivatives

In kinematics, there are a whole set of quantities that can describe an object's trajectory, e.g., the acceleration, which is defined to be the rate of change of velocity w.r.t. time:

$$\mathbf{a}(t) = \frac{\mathrm{d}\mathbf{v}(t)}{\mathrm{d}t}.$$

But the $\mathbf{v}$ is also the rate of change of position w.r.t. time. Thus,

$$\mathbf{a}(t) = \frac{\mathrm{d}\mathbf{v}(t)}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathrm{d}\mathbf{r}(t)}{\mathrm{d}t}\right) = \frac{\mathrm{d}^2\mathbf{r}(t)}{\mathrm{d}t^2},$$

where $\mathbf{r}$ is any position vector.

The $\mathrm{d}^2\mathbf{r}$ and $\mathrm{d}t^2$ is just a matter of symbolic manipulation and should only be interpreted as just a shorthand. $\mathbf{a}$ is called the **second order derivative** of $\mathbf{r}$ because you've differentiated $\mathbf{r}$ twice. **Higher order derivatives** of position w.r.t. time is listed in table 1.1.

Order	Name
1	Velocity/Speed
2	Acceleration
3	Jerk
4	Snap/Jounce
5	Crackle
6	Pop

TABLE 1.1 | HIGHER ORDER DERIVATIVES OF POSITION W.R.T. TIME

1.7 Expressing nature: basic differential equations

The dynamics of a physical system is universally described by the famous Newton's second law $\mathbf{F} = m\mathbf{a}(t)$, which in derivatives form becomes:

$$\mathbf{F} = m\frac{\mathrm{d}^2\mathbf{r}(t)}{\mathrm{d}t^2}. \tag{1.11}$$

Let's try to describe a simple system with this. In fig. 1.3, a ball is dropped from height h. The Earth's gravity pulls the ball with force mg where m is the mass of the

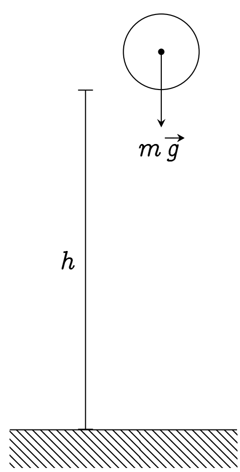


FIG. 1.3 | A BALL DROPPED FROM HEIGHT h

ball, and **g**, the acceleration from Earth's gravity. Newton's second law tells us that

$$\mathbf{F} = m\frac{\mathrm{d}^2\mathbf{r}(t)}{\mathrm{d}t^2} \tag{1.12}$$

$$m\mathbf{g} = m\frac{\mathrm{d}^2\mathbf{r}(t)}{\mathrm{d}t^2} \quad \text{or,} \quad \mathbf{g} = \frac{\mathrm{d}^2\mathbf{r}(t)}{\mathrm{d}t^2}, \tag{1.13}$$

which is sometimes called the **equation of motion**, which packs every information about this system you'd ever want. It directly reads as:

> The acceleration of the ball is equals to g.

If you've stayed for this long, congratulations! You now have the power to describe every physical systems with mathematical equations using derivatives: the measure of the rate of change. But it might not be as useful as yet, just as a hammer may seem useless if used to paint, derivatives falls apart when you ask about the future of the system. E.g., how long the ball takes to reach the ground? Or, what's the position of the ball at a certain time? That's the job of the integral to solve, and we'll do so in the next chapter.

1.8 Conclusion for Chapter 1

1. The concept of approaching can be used to bypass dividing by zero.
2. Derivatives are rate of change of a function w.r.t. a variable which can be evaluated by the method of increments.
3. Derivatives can be thought as the slope of a graph, or the tangent to a curve.
4. Physical systems can be described by differential equations of different forms. One of them is the Newton's formulation stated in eq. (1.11)

Remarks on chapter 1.1

1. In section 1.2, we zoomed in on the graph to approximate the function as a line. Actually, this is quite literally the whole idea of derivatives. If we dig in further in calculus, sometimes the rate of change analogy doesn't even make sense. However, saying that the derivative tries to approximate every function as a line works in all scenario. Though, it's quite abstracted away from the world.

CHAPTER 2

Integrals and antiderivative

Prerequisites: *chapter 1, sigma summation notation*

Terminologies: Every line, including straight, is a **curve**. You might see other textbooks use the term "*area under the graph*" to refer to integrals. But, a **graph** is *a diagram consisting of a line or lines, showing how two or more sets of numbers are related to each other* [4], not the curve itself. Therefore, I'll refer to a curve as any line that connects two points, whether straight or not. Now let's start.

2.1 Invitation to an impossible mission

2.1.1 The mindset of integral calculus

Literally everything that involves a quantity changing is written in the language of derivatives. When derivatives are used in equations to describe a system, it becomes *a differential equation*. Its solution contains the future of the entire system in question. However, it's quite hard to solve; some even impossible.

In classical physics, a **state** represents the configuration that the system *at one point in time*. To predict the system, we need to know the initial state of a system. Classical physics says that if you know the rules that the system plays by (In this case, $\mathbf{F} = m\frac{\mathrm{d}^2\mathbf{r}}{\mathrm{d}t^2}$), and the initial condition, we can always determine the state of the system at any point in time. That is, classical physics guarantees that there is always a function, which takes in time as input, and outputs the state of a system. And in order to predict the system, we must know that function. But what may that function be?

Consider the example from chapter 1. The equation of motion of the ball dropped from a height h is a second-order differential equation

$$g = \frac{\mathrm{d}^2\mathbf{r}(t)}{\mathrm{d}t^2}, \tag{2.1}$$

with initial condition being $\mathbf{r} = h$. The function that describes the state is $\mathbf{r}(t)$, which outputs the position of the ball at a certain time t. You can see that this function is the derivative sign. To solve the differential equation for this function, we have to isolate it out, and undo the derivative sign. But how?

It seems impossible at first. We only know that $\mathbf{r}(t)$ must satisfy eq. (2.1), i.e., the second derivative of $\mathbf{r}(t)$ is g. It's like we have to search through a gigantic pool of functions that mathematics have to offer to find a single function $\mathbf{r}(t)$ that satisfies eq. (2.1). It'd be like finding a needle in the haystack!

Of course, this is a textbook, there must be a solution. If there is a will, there is a way. You might have to reverse engineer derivatives, which might look tedious at first. But well, you might find something interesting along the way.

2.1.2 Brief notation of integral calculus

The $\int$, a.k.a. the **integral**[1], means to sum. This integral symbol is basically sigma summation symbol, but for infinitesimals. Therefore, we have bounds called **integral bounds**. E.g., if we sum a lot of little time step $\mathrm{d}t$ together from t_A to t_B we get $t_B - t_A$, the total time step. Thus,

$$\int_{t_A}^{t_B} \mathrm{d}t = t_B - t_A. \tag{2.2}$$

2.2 Finding a function in the haystack

Equation (2.1) has a second order derivative, let's go slowly and undo one derivative at a time. We'll undo the derivative of the R.H.S. to get the velocity first. Then, we'll undo the derivative again to get the position.

2.2.1 Step one: the velocity function from acceleration

One good strategy for solving any kind of equation is **separation of variables**. We isolate the variable we're interested in solving, which is $\mathbf{r}$, and move everything else to the other side. Here, write $\frac{\mathrm{d}^2\mathbf{r}(t)}{\mathrm{d}t^2}$ as $\frac{\mathrm{d}\mathbf{v}(t)}{\mathrm{d}t}$. Then, isolate $\mathbf{v}$ on one side and move $\mathbf{t}$ to the other.

$$\mathbf{g} = \frac{\mathrm{d}\mathbf{v}(t)}{\mathrm{d}t} \tag{2.3}$$

$$\mathrm{d}\mathbf{v}(t) = \mathbf{g}\,\mathrm{d}t. \tag{2.4}$$

This equation reads

> *A small change in velocity* $\mathrm{d}\mathbf{v}$ *is product of* $\mathbf{g}$ *and a small time interval* $\mathrm{d}t$.

[1] Which also looks like a beansprout

To find the total change in velocity, we sum up a lot of small changes in velocity $\mathrm{d}\mathbf{v}(t)$, which is equal to $\mathbf{g}\,\mathrm{d}t$. Because $\mathrm{d}v$ is directly proportional to $\mathrm{d}t$, the total change in $\mathbf{v}$ is simply $\mathbf{g}t$. However, *changes* doesn't say anything about the initial condition, so we add a term C to compensate. Therefore,

$$\mathbf{v}(t) = \mathbf{g}t + C. \tag{2.5}$$

To find what C is, just plug in the initial condition. If $\mathbf{v}(t = 0) = \mathbf{v}_0$, then

$$\mathbf{v}(t = 0) = v_0 = \mathbf{g} \times 0 + C \tag{2.6}$$

$$v_0 = C. \tag{2.7}$$

Thus,

$$\mathbf{v}(t) = \mathbf{g}t + \mathbf{v}_0. \tag{2.8}$$

We can also express these ideas symbolically using the integral symbol (section 2.1.2) as

$$\mathbf{g} = \frac{\mathrm{d}\mathbf{v}(t)}{\mathrm{d}t} \tag{2.9}$$

$$\mathrm{d}\mathbf{v} = \mathbf{g}\,\mathrm{d}t \tag{2.10}$$

$$\int \mathrm{d}\mathbf{v} = \int \mathbf{g}\,\mathrm{d}t \tag{2.11}$$

$$\mathbf{v} = \mathbf{g}t + \mathbf{v}_0. \tag{2.12}$$

2.2.2 Step two: the position function from velocity

Rewrite $\mathbf{v}$ as $\frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t}$, then do separation of variables.

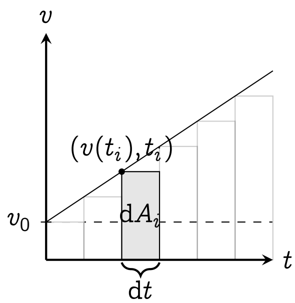
$$\mathbf{v} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \mathbf{g}t + \mathbf{v}_0 \tag{2.13}$$

$$\mathrm{d}\mathbf{r} = \mathrm{d}t(\mathbf{g}t + \mathbf{v}_0), \tag{2.14}$$

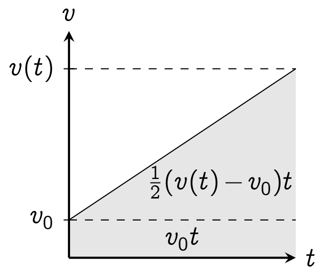
which reads,

> *A small change in position* $\mathrm{d}\mathbf{r}$ *is the product of* $\mathbf{g}t + \mathbf{v}_0$ *and a small time interval* $\mathrm{d}t$*.*

If we want to find the total change in position $\mathbf{r}$, we just have to sum all $\mathrm{d}\mathbf{r}$'s. This time, it's not as obvious, because there's $\mathbf{g}t + \mathbf{v}_0$, which is also changing with time. So at each point in time, the rate of change of position is different.



(A) A SUBDIVIDED GRAPH



(B) A GRAPH SHOWING THE AREA OF THE GRAY TRAPEZOID BROKEN INTO TWO PARTS

FIG. 2.1 | A v-t GRAPH OF A BALL DROPPED FROM A BUILDING

A good strategy in math if you don't know what to do is to just graph the function. The graph of eq. (2.13) is shown in fig. 2.1a. A small time interval represents a little step in the t axis. The curve shown in the graph represents $\mathbf{g}t + \mathbf{v}_0$. Therefore, $dt(\mathbf{g}t + \mathbf{v}_0)$ would just represent an area of a little rectangle dA as shown in the figure.

The total change in position is the sum all those rectangles. When $dt \to 0$, the sum of all $dt(\mathbf{g}t + \mathbf{v}_0)$ approaches the area under the graph, which can be calculated geometrically as shown in fig. 2.1b. Thus,

$$\mathbf{r} = \frac{1}{2}(\mathbf{v}(t) - \mathbf{t}_0)t + \mathbf{v}_0 t + C \tag{2.15}$$

$$= \frac{1}{2}(\mathbf{g}t - \mathbf{t}_0)t + \mathbf{v}_0 t + C \tag{2.16}$$

$$= \frac{1}{2}\mathbf{g} \times t^2 + \mathbf{v}_0 t + C \tag{2.17}$$

When $t = 0$, $\mathbf{r} = \mathbf{r}_0$. Therefore,

$$\mathbf{r}_0 = \frac{1}{2}\mathbf{g} \times 0^2 + \mathbf{v}_0 \times 0 + C$$

$$\mathbf{r}_0 = C.$$

Therefore,

$$\mathbf{r}(t) = \frac{1}{2}\mathbf{g}t^2 + \mathbf{v}_0 t + \mathbf{r}_0. \tag{2.18}$$

To model the trajectory of the ball, we set

1. $\mathbf{v}_0 = 0$ (object is dropped and starts at zero speed)
2. $\mathbf{r}_0 = 0$ (convenient initial condition placement)

thus we get,

$$\mathbf{r} = \frac{1}{2}\mathbf{g}t^2 \implies t = \sqrt{\frac{2\mathbf{r}}{\mathbf{g}}} \tag{2.19}$$

Since the ground is at $\mathbf{r} = h$, the time that the ball hits the ground is then $\sqrt{2h/\mathbf{g}}$.

2.2.3 Conclusion: area under the curve and antiderivative

The form of the differential equation that we've solved is in the form $\frac{\mathrm{d}x}{\mathrm{d}t} = f(x)$. We have to undo the derivative, and the simplest way is to do separation of variables and turn the equation into

$$\mathrm{d}x = f(x)\,\mathrm{d}t, \tag{2.20}$$

which reads,

> A small change in x, i.e., $\mathrm{d}x$ is represented by the area of a rectangle width $\mathrm{d}t$ and height $f(x)$.

And, the total change x is represented by the sum all those little rectangles, which is the area under the curve $f(x)$. Then, we add a constant C to compensate for the initial condition. In symbolic form introduced in section 2.1.2, it's just

$$\int \mathrm{d}x = \int g(x)\,\mathrm{d}t \tag{2.21}$$
$$x = \int g(x)\,\mathrm{d}t. \tag{2.22}$$

For now, we could say that integration is the reverse of derivatives. But to clearly see how this is linked for every function, we must study the fundamental theorem of calculus, which is the bridge between integration and differentiation.

Before we go there, let me clarify some terminologies. An **integral** refers to the area under the curve evaluated between two points. We say that an integral must have an **integral bound**. If we want to find the area under a function $f(x)$ from $x = a$ to $x = b$, we write it as

$$A = \int_a^b f(x)\,\mathrm{d}x. \tag{2.23}$$

This just reads

> The area A under the curve $f(x)$ from $x = a$ to $x = b$ is equal to the sum of the area of many thin stripes width $\mathrm{d}x$ height $f(x)$ that lies between $x = a$ and $x = b$.

The **antiderivative** however, refers to the function which takes in a value a, and output the integral of $f(x)$, evaluated from 0 to a. Therefore, if $A(a)$ is the antiderivative of $f(x)$, then

$$A(a) = \int_0^a f(x)\,\mathrm{d}x. \tag{2.24}$$

It also implies that if the area between a and b can be evaluated by

$$\int_a^b f(x)\,\mathrm{d}x = A(b) - A(a). \tag{2.25}$$

2.3 The fundamental theorem of calculus

The intuition of fundamental theorem of calculus states that derivatives and integrals are essentially inverse of each other. In this section, I'll clarify this fact and make it more rigorous.

Theorem 1: The (first) fundamental theorem of calculus If a function $f(x)$ has an antiderivative $A(x)$, then

$$\frac{\mathrm{d}A(x)}{\mathrm{d}x} = f(x). \tag{2.26}$$

If $A(x)$ is the antiderivative of $f(x)$, then

$$\int_0^x f(x)\,\mathrm{d}x = A(x). \tag{2.27}$$

The *actual* area of one of the stripes (not rectangles) width $\mathrm{d}x$ shown in fig. 2.2, it's obviously $A(x+\mathrm{d}x) - A(x)$. The riemann sum approximation approximates the area by a small rectangle area $f(x)\,\mathrm{d}x$. We can write the relation between the actual and the approximated area as

$$A(x+\mathrm{d}x) - A(x) \approx f(x)\,\mathrm{d}x. \tag{2.28}$$

To turn this into an equality, we add a correction term ϵ

$$A(x+\mathrm{d}x) - A(x) = f(x)\,\mathrm{d}x + \epsilon. \tag{2.29}$$

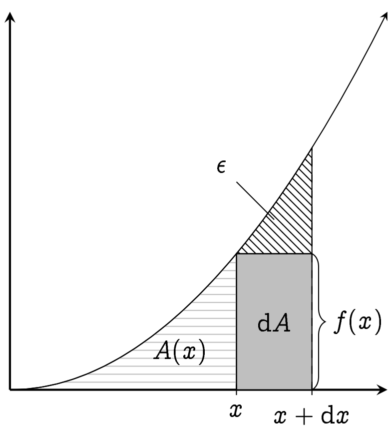


FIG. 2.2 | THE GEOMETRICAL INTERPRETATION OF PART ONE OF THE FUNDAMENTAL THEOREM OF CALCULUS (THEOREM 1).

If we let $\mathrm{d}x \to 0$, ϵ is negligible compared to $f(x)\,\mathrm{d}x$; therefore,

$$\lim_{\mathrm{d}x\to 0} A(x+\mathrm{d}x) - A(x) = \lim_{\mathrm{d}x\to 0} f(x)\,\mathrm{d}x. \tag{2.30}$$

Since $f(x)$ is not a variable that's controlled by the limit sign,

$$\begin{aligned}\lim_{\mathrm{d}x\to 0} A(x+\mathrm{d}x) - A(x) &= f(x) \lim_{\mathrm{d}x\to 0} \mathrm{d}x \\ \lim_{\mathrm{d}x\to 0} \frac{A(x+\mathrm{d}x) - A(x)}{\mathrm{d}x} &= f(x).\end{aligned}$$

And here, we see that the L.H.S. is just the derivative of $A(x)$ w.r.t. x, thus

$$\frac{\mathrm{d}A(x)}{\mathrm{d}x} = f(x), \tag{2.31}$$

or "*the rate of change in area is the function itself*". But the area function is given by the integral. This means

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\int f(x)\,\mathrm{d}x\right) = f(x): \tag{2.32}$$

integrals and derivatives are inverses of each other. If we rephrase theorem 1, we see that "*the integral is the cumulative effect of the function.*"

The second fundamental theorem of calculus,

> **Theorem 2: The second fundamental theorem of calculus (Newton-Leibniz rule)** If a function $f(x)$ has an antiderivative $A(x)$, then its indefinite integral from a to b is
>
> $$\int_a^b f(x)\,\mathrm{d}x = \int_0^b f(x)\,\mathrm{d}x - \int_0^a f(x)\,\mathrm{d}x = A(b) - A(a). \tag{2.33}$$

follows as a direct consequence of the geometrical interpretation of integrals: area under the curve.

The fundamental theorem of calculus allows us to evaluate integrals using derivatives. For example, if the derivative of x^2 is $2x$, then we also know that the integral of $2x\,\mathrm{d}x$ is just x^2, which we'll discuss how to do that in the next chapter.

2.4 How to calculate an integral? Riemann sum

In section 2.2, we used geometry to find the area under a curve. However, that is not always possible, e.g., try integrating fig. 2.3 geometrically[2]. It'd be impossible. But we can still approximate its area by slicing the area under the curve into thin rectangular stripes, then summing them. The approximated area is called the **Riemann sum**. Due to its computational cost, you don't really want to use this method. However, to develop a good intuition at the integral, we should still know its symbolic form.

Let there be a function $A(s)$ that represents the actual area of a function $f(x)$ from 0 to s. The approximated area is then

$$\sum_{i=0}^{k} \mathrm{d}A_i = \sum_{i=0}^{k} \text{width} \times \text{height} \tag{2.34}$$

[2]This is what you'd get if you solve the simple harmonic oscillator

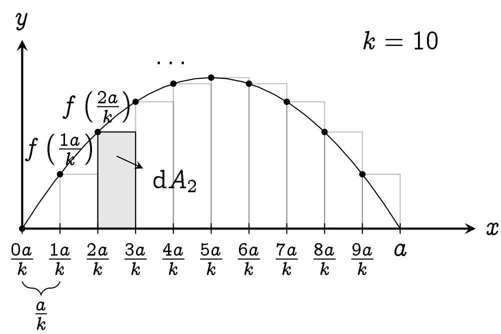


FIG. 2.3 | ILLUSTRATION OF RIEMANN SUM OF A FUNCTION $f(x)$ FROM 0 TO a BY SETTING $k = 10$

where k is the amount of subdivisions. From fig. 2.3, the width of each stripe is a / k, and the height of the i'th stripe is $f(ia / k)$. Therefore,

$$\sum_{i=0}^{k} \frac{a}{k} \times f\left(\frac{ia}{k}\right).$$

In fig. 2.3, we $k = 10$. The area of the second rectangle $\mathrm{d}A_2$ is

$$\mathrm{d}A_2 = \frac{a}{k} f\left(\frac{2a}{k}\right).$$

For an arbitrarily finite k, the Riemann sum is just an approximation. If you want to find the *actual* area under the curve, let $k \to \infty$. The limit as $k \to \infty$ is what we actually call the **integral**. Thus, we say

Definition 2: Naive definition of integrals

The (definite) integral, or the area under the curve of $f(x) = y$ from 0 to a, is defined as

$$\int_0^a \mathrm{d}A = \int_0^a f(x)\,\mathrm{d}x = \lim_{k\to\infty} \sum_{i=0}^{k} \frac{a}{k} \times f\left(\frac{ia}{k}\right) = A(a). \qquad (2.35)$$

where $\int$ is the integral sign[a]. Here, 0 is the lower bound of integration, and a, the higher bound. The function $A(x)$ is called the **antiderivative**, or the **indefinite integral** of $f(x)$.

[a]Famously known for looking like a beansprout

I shall put these definitions into perspective in the next two examples. It might use a bit of series knowledge. If you don't know, you can simply search up the summation identities that I'll use in *Wikipedia* [5].

Example 2.4.1: Riemann sum and antiderivative of x^2

Let $f(x) = x^2$, and let $A(a)$ be the antiderivative of $f(x)$, i.e., $A(a)$ is the area under the curve of $f(x)$ from 0 to a. Note that

$$\sum_{i=0}^{k} i^2 = \frac{k(k+1)(2k+1)}{6}, \tag{2.36}$$

The Riemann sum is then

$$\sum_{i=0}^{k} \frac{a}{k} \times f\left(\frac{ia}{k}\right) = \sum_{i=0}^{k} \frac{a}{k} \times \left(\frac{ia}{k}\right)^2 \tag{2.37}$$

$$= \sum_{i=0}^{k} \frac{a^3}{k^3} \times i^2 \tag{2.38}$$

$$= \frac{a^3}{k^3} \times \frac{k(k+1)(2k+1)}{6}. \tag{2.39}$$

To find the antiderivative, let $k \to \infty$. Notice that the $+1$ in the parenthesis are negligible when $k \to \infty$. Therefore, we can write the antiderivative as

$$A(a) = \int_0^a f(x)\,\mathrm{d}x \tag{2.40}$$

$$= \lim_{k\to\infty} \frac{a^3}{k^3} \times \frac{k(k+1)(2k+1)}{6} \tag{2.41}$$

$$= \lim_{k\to\infty} \frac{a^3}{k^3} \times \frac{2k^3}{6} \tag{2.42}$$

$$= \lim_{k\to\infty} \frac{x^3}{3} = \frac{x^3}{3}. \tag{2.43}$$

Example 2.4.2: Riemann sum and antiderivative of x^3

Let $f(x) = 4x^3$, and let $A(a)$ be the antiderivative of $f(x)$, i.e., $A(a)$ is the area under the curve of $f(x)$ from 0 to a. Note that

$$\sum_{i=0}^{k} i^3 = \left(\frac{k(k+1)}{2}\right)^2, \tag{2.44}$$

The Riemann sum is then

$$\begin{aligned}
\sum_{i=0}^{k} \frac{a}{k} \times f\left(\frac{ia}{k}\right) &= \sum_{i=0}^{k} \frac{a}{k} \times 4\left(\frac{ia}{k}\right)^3 \\
&= 4 \sum_{i=0}^{k} \left(\frac{x}{k}\right)^4 i^3 \\
&= 4 \left(\frac{x}{k}\right)^4 \left(\frac{k(k+1)}{2}\right)^2 .
\end{aligned}$$

To find the antiderivative, let $k \to \infty$. The $+1$ in $(k+1)$ can be ignored as $k \to \infty$. Therefore,

$$\begin{aligned}
A(a) &= \int_0^a f(x)\,\mathrm{d}x \\
&= \lim_{k\to\infty} 4\left(\frac{x}{k}\right)^4 \left(\frac{k^2}{2}\right)^2 \\
&= \lim_{k\to\infty} x^4 = x^4.
\end{aligned}$$

2.5 Basic applications of the integrals

2.5.1 Five equations of linear motion

So far, we have developed an intuition for derivatives, integrals, and their relationship. Let's apply those concepts to derive the famous equations of linear motion:

$$v(t) = v_0 + at, \tag{2.45}$$

$$s = \frac{1}{2}(v_0 + v(t))t, \tag{2.46}$$

$$s = v_0(t)t + \frac{1}{2}at^2, \tag{2.47}$$

$$s = v(t)t - \frac{1}{2}at^2, \tag{2.48}$$

$$v(t)^2 = v_0^2 + 2as. \tag{2.49}$$

Here, v_0 represents the initial velocity, $v(t)$, the velocity at any time t, a, the acceleration, and s, the displacement. These equations are derived from the Newton's second law on *constant/uniformed acceleration* motion in one dimension, which we can evaluate using the geometrical interpretation of integrals developed earlier.

Because we've constrained the object to accelerate at a, the force exerted must be ma. Newton's second law $F = m\frac{\mathrm{d}^2x}{\mathrm{d}t^2}$ then simplifies to $a = \frac{\mathrm{d}^2x}{\mathrm{d}t^2}$. The same idea applies, we rewrite the equation in terms of v.

$$a = \frac{\mathrm{d}v}{\mathrm{d}t} \tag{2.50}$$

Which reads, "*the rate of change of v w.r.t. t is a*" or "*the slope of the v-t curve is always equals to a*". Because a is constant, the v-t curve must be a straight line with slope a, shown in fig. 2.4.

Derivation of eq. (2.45). We take advantage of the linearness of the curve. Just pick two points on fig. 2.4, as already shown, then

$$m = a = \frac{\Delta v}{\Delta t} = \frac{v(t) - v_0}{t - 0}$$
$$a = \frac{v(t) - v_0}{t}$$

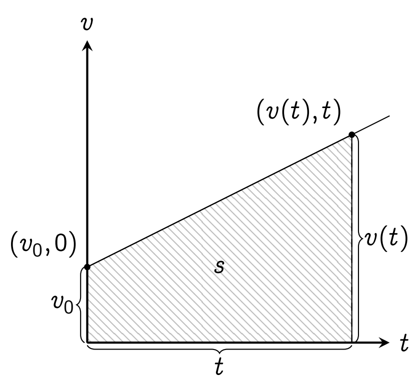


FIG. 2.4 | v-t GRAPH OF AN OBJECT UNDER CONSTANT OR UNIFORMED ACCELERATION

$$v(t) = v_0 + at. \qquad \square$$

Derivation of eq. (2.46). Use the reverse of theorem 1. We know that[3] $v = \frac{\mathrm{d}}{\mathrm{d}t}s$; therefore,

$$\int v \, \mathrm{d}t = s,$$

which reads "*The displacement s is the area under the curve of a v-t graph*". From fig. 2.4, the area under the curve is a trapezoid with side length v_0, $v(t)$, and width t. Thus,

$$s = \frac{1}{2}(v_0 + v(t))t,$$

which is just eq. (2.46): the area of a trapezoid. $\square$

Derivation of eqs. (2.47) to (2.49). We can arrange eq. (2.45) into three different ways, then plug in eq. (2.46).

First, $v(t) = v_0 + at$

$$\begin{aligned} s &= \frac{1}{2}(v_0 + v_0 + at)t \\ &= v_0 t + \frac{1}{2}at^2. \end{aligned}$$

Second, $v_0 = v(t) - at$

$$\begin{aligned} s &= \frac{1}{2}(v(t) - at + v(t))t \\ &= v(t)t - \frac{1}{2}at^2. \end{aligned}$$

Third, $t = \dfrac{v(t) - v_0}{a}$

$$\begin{aligned} s &= \frac{1}{2}(v(t) + v_0)\frac{(v(t) - v_0)}{a} \\ 2as &= (v(t) + v_0)(v(t) - v_0) \\ 2as &= v(t)^2 - v_0^2 \\ v(t)^2 &= 2as + v_0^2. \end{aligned} \qquad \square$$

2.5.2 The area of a circle

The perimeter of the circle is $2\pi r$. For now, we only know how to find areas of polygons, but we want to know the circle's area. So, is there any way to turn a circle into a polygon? From section 2.4, that the riemann sum can approximate areas under the curve using rectangular stripes. It'd be great if this circle can be turned into multiple rectangular stripes on a graph right?

So, we dissect a circle radially into small rings $\mathrm{d}r$ thin, as shown

[3] Here, x is replaced with s to represent displacement in one dimension.

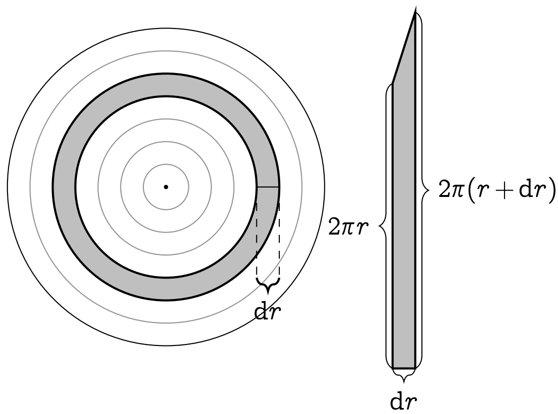


FIG. 2.5 | (LEFT) DISSECTING A CIRCLE RADIUS R RADIALLY INTO RINGS, EACH RING $\mathrm{d}r$ THIN. (RIGHT) STRETCHING A RING INTO A TRAPEZOID (NOT TO SCALE).

in fig. 2.5. Then, stretch all the rings into a very thin trapezoid-like shape. It might seem impossible at first, but considering that our ring is very thin, it's quite easy to stretch it without breaking. I encourage you to grab a piece of paper, cut a really thin ring and try it out. If you actually do it, it'll look a bit warped. However, the warpedness will go away the thinner you go.

Since we sliced our trapezoid from a circle, for a stripe positioned at r, the inner side will be $2\pi r$ long and the outer side, $2\pi(r + dr)$. The area of the little trapezoid $\mathrm{d}A$ then becomes

$$\begin{aligned}\mathrm{d}A &= \frac{1}{2}\,\mathrm{d}r(2\pi r + 2\pi(r + dr)) \\ &= 2\pi r\,\mathrm{d}r + 2\pi\,\mathrm{d}r^2.\end{aligned}$$

As $\mathrm{d}r \to 0$, $\mathrm{d}r^2$ becomes negligible. Therefore,

$$\mathrm{d}A = 2\pi r\,\mathrm{d}r. \tag{2.51}$$

This equation says that for $\mathrm{d}r \to 0$, the trapezoid becomes a rectangle side length $\mathrm{d}r$ and $2\pi r$. Now, we have to sum it together. If we can put all these rectangles onto a graph, we can easily use the Riemann's sum to evaluate it. A natural way to do this is to put all the rectangles that we got from stretching the rings of the circle onto a graph one by one. The result would look something like fig. 2.6[4].

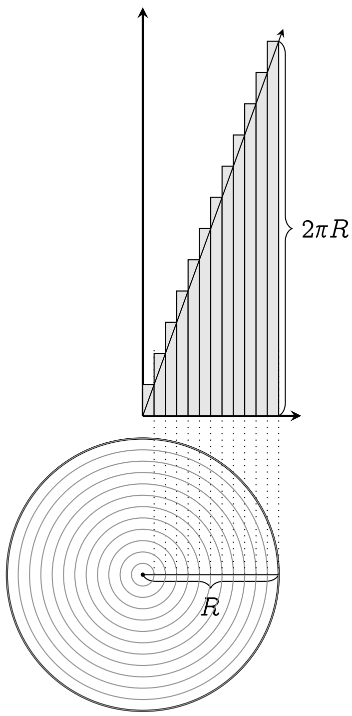


FIG. 2.6 | REARRANGING ALL THE APPROXIMATED RECTANGLES ONTO A GRAPH (NOT TO SCALE).

For every stripe at r, its height is $2\pi r$. If you were to plot the height of all rectangles when $\mathrm{d}r \to 0$, it'll eventually look like a curve that's given by $f(r) = 2\pi r$. We're interested in the area of the circle from 0 to R. Now, it's transformed into the area under the curve of $f(r) = 2\pi r$: a triangle with base R and height $f(R) = 2\pi R$. Thus, the area of a circle becomes

$$A = \frac{1}{2}R(2\pi R) = \pi R^2.$$

And there you go, you've turned essentially turned circle into a triangle and evaluate its area from there. I'd like to end off this chapter by mentioning the spirit of mathematics. Sometimes, you can't solve the problem directly. Most of the time, you have to re-frame the problem into another more-solvable problem. Problems like these often have the most sublime connections to the foundations of mathematics. This is a common

[4] Not to scale

theme in most of mathematics, especailly calculus. So, be sure to keep this in mind while reading through.

2.6 Conclusion for Chapter 2

1. Riemann sum are used to approximate areas under the curve of a function by using little rectangles then summing it.
2. Integrals or anti-derivatives are functions that output the area under the graph of other functions.
3. The limit where the width of the rectangles in the Riemann sum approaches zero, the Riemann sum becomes an integral.
4. Integrals are the cumulative effect of a function.
5. Integrals and derivatives are inverses of each other, and they're related by the fundamental theorem of calculus
6. Integrals can be used in various ways by reframing questions into another simpler question.

CHAPTER 3

Basic derivatives and antiderivatives

This chapter covers the derivatives and integrals of common functions: polynomials, exponential, and logarithms; focusing on their geometrical interpretation.

Prerequisites: *binomial theorem (appendix B), basic trigonometry, derivatives, and integrals*

Terminologies: The integral refers to the area under the curve. The antiderivative refers to the function $A(x')$ that outputs the area under the curve from 0 to x'.

3.1 Basic rules

The chain rule The Leibniz' notation treats derivative as fractions; thus, we can cancel terms (eq. (1.9)). We exploit this property to find derivatives of composite functions. Consider two functions $f(x)$ and $g(x)$, the derivative of $f(g(x))$ can be found by a simple substitution. First let $g(x) = u$.

$$\frac{\mathrm{d}}{\mathrm{d}x} f(g(x)) = \frac{\mathrm{d}f(u)}{\mathrm{d}x}.$$

Let $1 = \frac{\mathrm{d}u}{\mathrm{d}u}$, then

$$\frac{\mathrm{d}}{\mathrm{d}x} f(u) = \frac{\mathrm{d}}{\mathrm{d}x} f(u) \times 1 = \frac{\mathrm{d}f(u)}{\mathrm{d}x}\frac{\mathrm{d}u}{\mathrm{d}u}.$$

Perform a change of denominator, then substitute back $u = g(x)$:

$$\frac{\mathrm{d}}{\mathrm{d}x} f(g(x)) = \frac{\mathrm{d}f(u)}{\mathrm{d}u} \times \frac{\mathrm{d}g(x)}{\mathrm{d}x}.$$

This is what we call the chain rule, or more generally

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}y}{\mathrm{d}u} \times \frac{\mathrm{d}u}{\mathrm{d}x}. \tag{3.1}$$

where u is a function of y.

The chain rule holds the intuition of how rate of changes relate to each other. E.g., the cheetah's speed is 10 times the bicycle's speed, which is 4 times the walking speed. The ratio between the cheetah's speed compared to walking speed would obviously be $10 \cdot 4 = 40$:

$$\frac{\mathrm{dCheetah}}{\mathrm{dWalking}} = \frac{\mathrm{dCheetah}}{\mathrm{dBicycle}} \times \frac{\mathrm{dBicycle}}{\mathrm{dWalking}}. \tag{3.2}$$

Integral constant In section 2.2, each time we evaluate the antiderivative, we add an initial condition term, i.e., v_0 and r_0. So for any function $f(x)$,

$$\int f(x)\,\mathrm{d}x = A(x) + C \tag{3.3}$$

where C is any constant. However, theorem 2 still holds for integrals with bounds.

Always add an integral constant after evaluating the antiderivative

Integral of the infinitesimal The antiderivative of the small rectangles $\mathrm{d}x$ is just the total rectangle x plus the integral constant C.

$$\int \mathrm{d}x = x + C. \tag{3.4}$$

Rules of equality If two arguments are equal, their derivatives and antiderivatives w.r.t. the same variable must also be equal.

$$\text{If } f = g, \text{ then } \frac{\mathrm{d}f}{\mathrm{d}x} = \frac{\mathrm{d}g}{\mathrm{d}x}, \text{ and } \int f \,\mathrm{d}x = \int g \,\mathrm{d}x + C$$

Derivative of a constant A constant doesn't change; thus, the derivative of a constant is zero.

$$\frac{\mathrm{d}}{\mathrm{d}x}(c) = 0. \tag{3.5}$$

3.2 Linearity of differentiation and integration

A **linear** mathematical entity are defined as anything that is compatible with addition and scaling. E.g., for a function $f(x)$ to be linear, it must obey

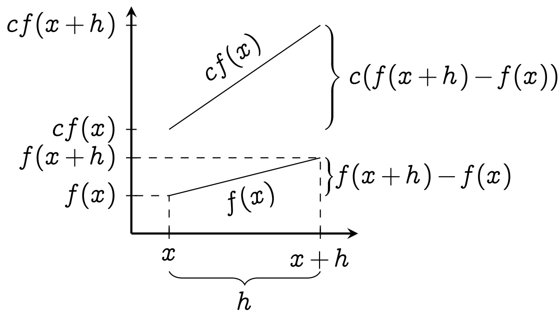
1. Additivity: $f(x + y) = f(x) + f(y)$
2. Homogeneity of degree one: $f(ax) = af(x)$ for all constant a

The simplest linear function there is, is a line that passes through the origin: $f(x) = ax$, in which

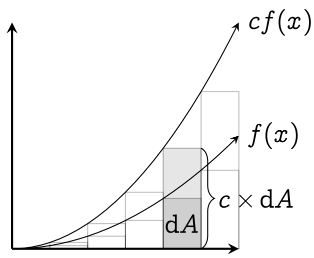
$$f(x + y) = a(x + y) = ax + ay = f(x) + f(y), \tag{3.6}$$

$$f(cx) = a(cx) = c(ax) = cf(x). \tag{3.7}$$

Because this property, we associate a line with being linear. However, linearity doesn't have to always refer to lines.



(A) FOR DERIVATIVES



(B) FOR INTEGRALS

FIG. 3.1 | CONSTANT MULTIPLE RULES

Both derivatives and integrals are linear; for any function $f(x)$, $g(x)$, and constants a, b, the following rules follow.

The constant multiple rule:

$$\frac{\mathrm{d}}{\mathrm{d}x}(af(x)) = a\frac{\mathrm{d}f(x)}{\mathrm{d}x}, \qquad \text{Illustrated in fig. 3.1a} \tag{3.8}$$

$$\int af(x)\,\mathrm{d}x = a\int f(x)\,\mathrm{d}x. \qquad \text{Illustrated in fig. 3.1b} \tag{3.9}$$

The geometrical interpretation of these two rules are very simple. When a function is multiplied by a constant c, its value is increased by a factor of c everywhere. The slope must also be increased by c, and thus the function's area also.

The sum rule:

$$\frac{\mathrm{d}}{\mathrm{d}x}(f(x)+g(x)) = \frac{\mathrm{d}}{\mathrm{d}x}f(x) + \frac{\mathrm{d}}{\mathrm{d}x}g(x), \tag{3.10}$$

$$\int f(x) + g(x)\,\mathrm{d}x = \int f(x)\,\mathrm{d}x + \int g(x)\,\mathrm{d}x. \tag{3.11}$$

I.e., adding two functions increases its slope, thus also increasing its area under the graph.

3.3 Derivatives and antiderivatives of polynomials

Now that we've discussed the "trivial rules", we're ready to tackle the easiest family of functions: the **polynomials**. They're in the form

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \dots . \tag{3.12}$$

Let's focus on the antiderivative first; we'll see later how the antiderivative of polynomials can be found with just a simple substitution trick.

The form mentioned in eq. (3.12) is quite useless if we want to make progress. We can break it down by using the linearity of derivatives:

$$\begin{aligned} \frac{\mathrm{d}f(x)}{\mathrm{d}x} &= \frac{\mathrm{d}}{\mathrm{d}x}(a_0 + a_1 x^1 + a_2 x^2 + a_3 x^3 + \dots) && (3.13) \\ &= a_1 \frac{\mathrm{d}x^1}{\mathrm{d}x} + a_2 \frac{\mathrm{d}x^2}{\mathrm{d}x} + a_3 \frac{\mathrm{d}x^3}{\mathrm{d}x} + \dots . && (3.14) \end{aligned}$$

Now we're left with derivatives of *monomials* in the form of x^n. So let's do that instead.

3.3.1 Derivatives of monomials: the power rule

The method of increments allow us to quickly evaluate the derivative of x^n.

$$\frac{\mathrm{d}}{\mathrm{d}x}(x^n) = \lim_{h\to 0} \frac{(x+h)^n - x^n}{h}. \tag{3.15}$$

By using the binomial expansion (appendix B),

$$= \lim_{h\to 0} \frac{\left(\sum_{k=0}^{n} \binom{n}{k} x^{n-k} \cdot h^n\right) - x^n}{h} \tag{3.16}$$

$$= \lim_{h\to 0} \frac{\binom{n}{0}x^n h^0 + \binom{n}{1}x^{n-1}h^1 + \binom{n}{2}x^{n-2}h^2 + \dots \binom{n}{n} + x^0h^n - x^n}{h} \tag{3.17}$$

$$= \lim_{h\to 0} \frac{x^n + nx^{n-1}h^1 + \binom{n}{2}x^{n-2}h^2 + \cdots + h^n - x^n}{h} \tag{3.18}$$

$$= \lim_{h\to 0} nx^{n-1} + \binom{n}{2}x^{n-2}h + \binom{n}{3}x^{n-3}h^2 + h^{n-1} \tag{3.19}$$

The terms with h vanishes when $h \to 0$; therefore,

$$\frac{\mathrm{d}}{\mathrm{d}x}(x^n) = nx^{n-1}, \tag{3.20}$$

which is what we call the power rule. Because the binomial theorem work for all real numbers, this is true for any powers of n. But what about the geometrical interpretation given?

Let's now focus on the geometrical interpretation of the derivative of x^2: a function that represents the area of a square sidelength x. Its derivative then represents the ratio between the change in x, and the change of area when the sidelength is increased by $\mathrm{d}x$.

Illustrated in fig. 3.2,

$$\frac{\mathrm{d}}{\mathrm{d}x}(x^2) = \lim_{\mathrm{d}x\to 0} \frac{x\,\mathrm{d}x + x\,\mathrm{d}x + \mathrm{d}x^2}{\mathrm{d}x}. \tag{3.21}$$

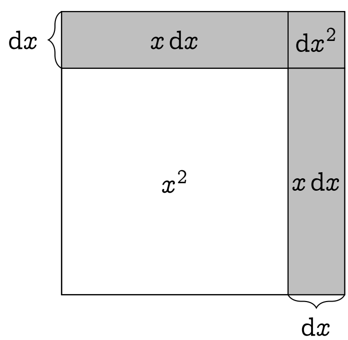


FIG. 3.2 | INTERPRETATION OF $\frac{\mathrm{d}x^2}{\mathrm{d}x}$ ($\mathrm{d}x$ EXAGGERATED)

When $\mathrm{d}x \to 0$, the $\mathrm{d}x^2$ square would just become a single point. Compared to the big $x\,\mathrm{d}x$ on the side, it's negligible; therefore, we can safely ignore it. The equation then becomes

$$\frac{\mathrm{d}}{\mathrm{d}x}(x^2) = \lim_{\mathrm{d}x\to 0} \frac{2x\,\mathrm{d}x}{\mathrm{d}x} = 2x,$$

which is equivalent to the result from the power rule.

Deriving $\frac{\mathrm{d}}{\mathrm{d}x}(x^3)$ geometrically shouldn't be too hard either. Take a cube sidelength x; increase its side length by $\mathrm{d}x$, and compute the ratio between the change in volume and $\mathrm{d}x$. The final answer should be $3x^2$. You could try with more higher order of n, but it'd be very hard to visualize.

Here I leave some exercises which shouldn't be too hard to do

1. $\frac{\mathrm{d}}{\mathrm{d}x}(x^2 - 2x + 16)$ $\qquad 2x - 2$
2. $\frac{\mathrm{d}}{\mathrm{d}x}(x^3 + x^2 + x + 1)$ $\qquad 3x^2 + 2x + 1$
3. $\frac{\mathrm{d}}{\mathrm{d}x}(3x^4 + 24x^3 - 2x^2 - 32x + 88)$ $\qquad 12x^3 + 72x^2 - 4x - 32$

3.3.2 Antiderivatives of polynomials: the reversed power rule

For the antiderivative of x^n, we use the fundamental theorem of calculus (theorem 1) on the power rule (eq. (3.20)),

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

Substitute n with $n+1$

$$\int (n+1)x^n\,\mathrm{d}x = x^{n+1} + C,$$

and by the linearity of integrations (eq. (3.9)), we get the **reversed power rule**:

$$\int x^n\,\mathrm{d}x = \frac{x^{n+1}}{n+1} + C. \tag{3.22}$$

Notice, this rule does not work for $n = 1$ because we can't divide by zero...*or can we*??? (Discussed further in section 3.8)

3.3.3 Extending the equations of linear motion

Let's use the power rule to extend the equation of linear motions in section 2.5.1 to include a constant jerk j, i.e.,

$$\frac{\mathrm{d}a}{\mathrm{d}t} = j \tag{3.23}$$

Then, we can move the $\mathrm{d}t$ around and integrate both sides by using the reversed power rule.

$$\begin{aligned}
\int j\,\mathrm{d}t &= \int \mathrm{d}a \\
j\int \mathrm{d}t &= \int \mathrm{d}a && \text{Linearity of integrals} \\
jt + a_0 &= a && \text{Integral constant } a_0
\end{aligned}$$

Because a is the derivative of v,

$$\begin{aligned}
\frac{\mathrm{d}v}{\mathrm{d}t} &= jt + a_0 \\
\int \mathrm{d}v &= \int jt + a_0\,\mathrm{d}t \\
v &= j\int t\,\mathrm{d}t + \int a_0\,\mathrm{d}t && \text{Linearity of integrals} \\
v &= \frac{1}{2}jt^2 + a_0 t + v_0. && \text{Reversed power rule}
\end{aligned}$$

We add v_0 as an integral constant in a similar manner. Since v is the derivative of r,

$$\begin{aligned}
\frac{\mathrm{d}r}{\mathrm{d}t} &= \frac{1}{2}jt^2 + a_0 t + v_0 \\
\int \mathrm{d}r &= \int \frac{1}{2}jt^2 + a_0 t + v_0\,\mathrm{d}t \\
r &= \frac{1}{2}j\int t^2\,\mathrm{d}t + a_0\int t\,\mathrm{d}t + v_0\int \mathrm{d}t
\end{aligned}$$

$$r = \frac{1}{6}jt^3 + \frac{1}{2}a_0t^2 + v_0t + r_0.$$

And there you have it! Further extensions can be made from here: the jerk could be time-dependent. But that's probably enough to illustrate my point. If you'd like to try, extend the equation of motion to have a constant snap: $\frac{\mathrm{d}j}{\mathrm{d}t} = s$. You should get

$$r = \frac{1}{24}st^4 + \frac{1}{6}jt^3 + \frac{1}{2}a_0t^2 + v_0t + r_0.$$

3.4 Exponentials and growth

t	$V(t)$	$V(t) - V(t-1)$
0	1	
1	2	$2 - 1 = 1$
2	4	$4 - 2 = 2$
3	8	$8 - 4 = 4$
4	16	$16 - 8 = 8$
5	32	$32 - 16 = 16$
6	64	$64 - 32 = 32$
7	128	$128 - 64 = 64$
8	256	$256 - 128 = 128$
9	512	$512 - 256 = 256$
10	1024	$1024 - 512 = 12$

TABLE 3.1 | TABLES OF 2^x PLOTTED AT INTERVAL 1 FROM 0 TO 10

The next function that we're going to discuss is exponentials: the mathematical representation of growth and decay. As an example, let's say there's a magical drop of water that doubles its volume V every hour. I.e., for any time t,

$$V(t+1) = 2V(t). \tag{3.24}$$

If the drop starts at one unit of volume, $V(0) = 1$; thus,

$$\begin{aligned} V(1) &= 2V(0) = 2, \\ V(2) &= 2V(1) = 2(2) = 2^2, \\ V(3) &= 2V(2) = 2(2^2) = 2^3, \\ V(4) &= 2V(3) = 2(2^3) = 2^4, \\ &\vdots \end{aligned}$$

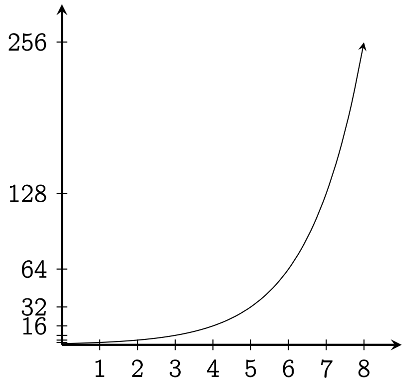


FIG. 3.3 | EXPONENTIAL FUNCTION 2^x PLOTTED FROM 0 TO 8

It's clear that the pattern is $V(t) = 2^t$: an exponential function. A natural question to ask is "*what is its rate of change?*". Let's start by plotting the function over time (fig. 3.3). But, exponentials grow too quick to plot! By $V(7)$, we're already in the hundreds. So It'd probably be better to list the values of each point on a table. Notice that on the right most column of table 3.1, the difference between $V(t)$ and $V(t-1)$ is exactly $V(t-1)$: the function changes as much as its past-self. So does that mean that $\frac{\mathrm{d}}{\mathrm{d}x}(2^x) = 2^x$?

Well sadly not, but close. See, table 3.1 only shows a *discrete* step. You can write it out as

$$\frac{2^{x+1} - 2^x}{1} = 2^x \left(\frac{2-1}{1}\right) = 2^x, \tag{3.25}$$

that is why $V(t) - V(t-1) = V(t-1)$. But we still need the method of increments to calculate the derivative of 2^x:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x}(2^x) &= \lim_{h\to 0} \frac{2^{x+h} - 2^x}{\mathrm{d}x} \\ &= 2^x \lim_{h\to 0} \left(\frac{2^h - 1}{h}\right). \end{aligned}$$

Now you could try plugging in a really small value of h, say 0.000001. The term $\frac{2^h - 1}{h}$ will approach 0.69314 … . If you try other bases of exponents, say 3, you might see a pattern emerging.

$$\frac{\mathrm{d}}{\mathrm{d}x}(3^x) = 3^x \lim_{h \to 0} \frac{3^h - 1}{h}. \tag{3.26}$$

The rate of change of an exponential function is always itself times a proportionality constant. For 3^x, it's about $1.09851\ldots$. If we could find a number n where $\frac{n^h - 1}{h} = 0$, we'd have a very pretty function which it is its own derivative. So let's find that!

3.4.1 A function that is its own derivative

Let's set a goal: find the function that is its own derivative. I shall introduce a substantial concept in calculus: the expansion of functions. Every function has a polynomial expansion[1] called the **power series**. For every $f(x)$,

$$f(x) = a_0 + a_1 x^1 + a_2 x^2 + a_3 x^3 + \ldots. \tag{3.27}$$

E.g., $\sin(x)$ can be written as

$$\sin(x) = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \frac{1}{7!}x^7 + \ldots. \tag{3.28}$$

we will derive this expression later in **??**. For now, we can just use eq. (3.27) to find the expression for the function that is its own derivative.

We've seen that the exponential is a possible candidate for a function that is its own derivative. Now, assume that for some real number n,

$$\frac{\mathrm{d}}{\mathrm{d}x}(n^x) = n^x. \tag{3.29}$$

Then, we use the polynomial expansion and the power rule,

$$\frac{\mathrm{d}}{\mathrm{d}x}(a_0 + a_1 x^1 + a_2 x^2 + a_3 x^3 + \ldots) = a_1 + 2a_2 x^1 + 3a_3 x^2 + 4a_4 x^3 + \ldots$$

[1] Although the convergence of the series derived is quite questionable; thankfully, the power series of n^x converges everywhere.

$$= n^x$$

If the function is its own derivative, the polynomial expansion of the function and its derivative must be the same.

$$n^x = a_0 + a_1 x^1 + a_2 x^2 + a_3 x^3 + \dots, \tag{3.30}$$

$$n^x = a_1 + 2a_2 x^1 + 3a_3 x^2 + 4a_4 x^3 + \dots. \tag{3.31}$$

Since both are polynomials, we can match the coefficient here:

$$\begin{aligned} a_0 &= a_1 \\ a_1 &= 2a_2 \\ a_2 &= 3a_3 \\ a_3 &= 4a_4 \\ &\vdots \end{aligned} \tag{3.32}$$

a_0 and a_1 is relatively easy to find. As we've seen, n must be between 2 and 3. By the properties of exponentials, $x = 0 \implies n^x = 1$. We can then plug $x = 0$ and set $n^x = 1$ into eq. (3.31):

$$\begin{aligned} 1 &= a_0 + a_1(0)^1 + a_2(0)^2 + a_3(0)^3 + \dots \\ a_0 &= 1. \end{aligned}$$

Since $a_0 = a_1$, a_1 must also be 1. We can then go back to eq. (3.32) and get

$$n^x = 1 + 1 + \frac{1}{2!}x + \frac{1}{3!}x^3 + \frac{1}{4!}x^4 + \dots$$

The pattern here is clearly $a_n = n!$. If we want to find n, we just let $x = 1$.

$$n = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \dots,$$

and there we have an expression for n which is an irrational number. If you work this out, it's around $2.71828\dots$. Because $(2.71828\dots)^x$ is its own derivative, it's very useful in mathematics and appears everywhere, even at the seams of mathematics that doesn't even seems related to growths: the patterns of prime number, this constant 2.71828 ... has a name and symbol: the Euler's number[2], written as e where

$$\mathrm{e} = \sum_{i=0}^{\infty} \frac{1}{i!} \tag{3.33}$$

3.4.2 Another interpretation of e: infinite bank interests

There are two types of bank interests: simple and compound. Simple interests is the thing that you don't really want: the interest is always the same and doesn't grow with your account. You can calculate it by using

$$n(t) = n_0 + tr \tag{3.34}$$

where $n(t)$ is the total money at time t, n_0 the initial money in your bank account, and r, the interest rate.

Compound interest in the other hand calculates your interest based on how much money you have at that moment:

$$n(t+1) = n(t)r + n(t). \tag{3.35}$$

We can find the expression for $n(t)$ in a similar fashion to what we've done in eq. (3.24). You'll get

$$n(t) = (1+r)^t n_0 \tag{3.36}$$

which is an exponential function.

Let's say you deposit 100$ into a bank and the bank is offering you two options on **compound interests** rate. 1) Take 100% interest in 1 year, 2) Take 100 / 2% twice a year, or 3) Take 100 / 356% daily. If you take option

[2] Not to be confused with the "Euler's constant" which is another constant written γ, and is around 0.57721 ...

one, you'd end up with 200\$. Option two takes you to 225\$, and option three takes you to around 271.447 ... \$. You might see a theme here. If you get 100 / $n\%$ interest, n times a year, the result keeps getting higher. Is there an upper limit to this?

If we write it in terms of limits as $n \to \infty$ and use eq. (3.36), the compound interests formula,

$$x = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n n_0.$$

Where x is the total money after a year. We're interested in the upper limit, so we'll just let n_0 for now. The expression will become

$$x = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n \tag{3.37}$$

Technically, we could go in and substitute a very high n, such as 1000000. But I believe you could already see that it would be a nightmare to calculate: exponentiation is not at all an easy task. However, notice that from option three earlier, the total money is 271.447 ... \$ which is suspiciously similar to e at 2.71828 ... . If eq. (3.37) equals eq. (3.32), we'd find the upper limit for this problem and solve the mystery.

We can use the binomial theorem on eq. (3.37) and get

$$\begin{aligned}
&\lim_{n\to\infty} \sum_{k=0}^{n} \binom{n}{k} 1^{n-k} \frac{1}{n^k} \\
&= \lim_{n\to\infty} \binom{n}{0}\frac{1}{n^0} + \binom{n}{1}\frac{1}{n^1} + \binom{n}{2}\frac{1}{n^2} + \binom{n}{3}\frac{1}{n^3} + \dots .
\end{aligned}$$

Then, we use the definition of n choose k,

$$= 1 + \lim_{n\to\infty} \frac{n!}{1!(n-1)!}\frac{1}{n^1} + \frac{n!}{2!(n-2)!}\frac{1}{n^2} + \frac{n!}{3!(n-3)!}\frac{1}{n^3} + \dots .$$

Now, we can cancel the $n!$ on the numerator to the denominator and isolate the factorials.

$$= 1 + \lim_{n\to\infty} \frac{n(n-1)!}{(n-1)!}\frac{1}{1!n^1} + \frac{n(n-1)(n-2)!}{(n-2)!}\frac{1}{2!n^2} + \dots$$

$$= 1 + \frac{1}{1!} + \lim_{n\to\infty} \frac{n(n-1)}{2!n^2} + \frac{n(n-1)(n-2)}{3!n^3} + \frac{n(n-1)(n-2)(n-3)}{4!n^4} + ...$$

Notice, as $n \to \infty$, the ratio between $n + R$ and n where R is any real numbers would be literally negligible. For every terms in our series, both the numerator and the denominator has the same polynomic degrees. Therefore, all the n's in the series cancel out and we get

$$x = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + ... \tag{3.38}$$

which is literally eq. (3.32); thus,

$$\mathrm{e} = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n. \tag{3.39}$$

In conclusion, the upper limit that the bank can give you is e, which should make sense geometrically. Because we're gradually turning a discrete interest into a continuous one, e should appear in the limit of the continuous bank interests.

The limit definition of e allows us to express e^x in terms of limits.

$$\mathrm{e}^x = \left(\lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n\right) \tag{3.40}$$

$$= \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^{nx} \tag{3.41}$$

If $n = u \,/\, x$, then

$$\mathrm{e}^x = \lim_{\frac{u}{x}\to\infty} \left(1 + \frac{u}{h}\right)^h \tag{3.42}$$

For $\dfrac{u}{x} \to \infty$ to be true for all finite x, u must approach ∞; therefore,

$$\mathrm{e}^x = \lim_{u\to\infty} \left(1 + \frac{u}{h}\right)^h \tag{3.43}$$

3.4.3 The antiderivative of exponential functions

It should be trivial that if e^x is the derivative of itself, so is its antiderivative

$$\int \mathrm{e}^x \,\mathrm{d}x = \mathrm{e}^x + C. \tag{3.44}$$

For other bases, we could use the fundamental theorem of calculus (theorem 1) to find its antiderivative. I.e., if

$$\frac{\mathrm{d}(n^x)}{\mathrm{d}x} = n^x \lim_{h\to 0} \frac{n^h - 1}{h},$$

then

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}(n^x) &= n^x \lim_{h\to 0} \frac{n^h - 1}{h} \\
n^x &= \lim_{h\to 0} \frac{n^h - 1}{h} \int n^x \,\mathrm{d}x \\
\int n^x \,\mathrm{d}x &= n^x \lim_{h\to 0} \left(\frac{n^h - 1}{h}\right)^{-1} + C.
\end{aligned}$$

The term in the limit sign still appears here. If we want to uncover the origin of this term, we must discuss the logarithms.

3.5 Logarithms

Logarithms are inverses of exponential, like how roots are inverses of monomials. To illustrate what I mean,

$$\text{For monomials, } \sqrt[a]{x^a} = x, \text{ but with logarithms, } \log_a(a^x) = x.$$

They have the following properties:

$$\log_a(x) + \log_a(y) = \log_a(xy), \tag{3.45}$$

$$\log_a(x) - \log_a(y) = \log_a\left(\frac{x}{y}\right), \tag{3.46}$$

$$\log_a(x) = \frac{\log_b(x)}{\log_b(a)}. \tag{3.47}$$

Since e^x is shown to be a very important function in modelling continuous growth, and is its own derivative, we give its inverse function its own name: the **natural logarithm**, written as $\ln(x)$.

What's the derivative of $\ln(x)$? By the method of increments,

$$\begin{aligned}
\frac{d}{dx}\ln(x) &= \lim_{h\to 0}\frac{\ln(x+h)-\ln(x)}{h} && (3.48)\\
&= \lim_{h\to 0}\frac{\ln(\frac{x+h}{x})}{h} && (3.49)\\
&= \lim_{h\to 0}\frac{\ln(1+\frac{h}{x})}{h} && (3.50)\\
&= \lim_{h\to 0}\ln\left(\left(1+\frac{h}{x}\right)^{\frac{1}{h}}\right). && (3.51)
\end{aligned}$$

This form is very similar to the limit definition of the exponential function e^x, but with some terms swapped around. So let's try to fit this limit to the form of e^x. I shall let $h = \frac{1}{n}$. If $h \to 0$, then $n \to \infty$. Therefore,

$$\begin{aligned}
\frac{d}{dx}\ln(x) &= \lim_{n\to\infty}\ln\left(\left(1+\frac{1}{nx}\right)^{n}\right) && (3.52)\\
&= \lim_{n\to\infty}\ln\left(\left(1+\frac{1/x}{n}\right)^{n}\right). && (3.53)
\end{aligned}$$

Recall that

$$e^x = \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n. \tag{3.54}$$

Thus,

$$\lim_{n\to\infty}\left(1+\frac{1/x}{n}\right)^n = e^{\frac{1}{x}}. \tag{3.55}$$

which fits the form of eq. (3.53). Therefore,

$$\begin{aligned}
\frac{d}{dx}\ln(x) &= \lim_{n\to\infty}\ln(e^{\frac{1}{x}}) && (3.56)\\
&= \lim_{n\to\infty}\frac{1}{x} = \frac{1}{x}: && (3.57)
\end{aligned}$$

the derivative of the natural logarithm of x is the reciprocal of x.

This result can also be interpretted geometrically. If the exponential grows very quickly when x increases, its inverse must grow slow. The slow growing rate is directly captured by the function $\frac{1}{x}$.

Here is the perfect time that I'll give you a glimpse into calculus with many variables: multivariable calculus. I shall present another way to derive the derivative of $\ln(x)$: **implicit differentiation**.

Functions in the form $f(x) = y$ are called **explicit functions**: the relations between the two variables are written explicitly. Functions that can't be written as $f(x) = y$ are called **implicit functions**, e.g., $x^2 + y^2 = r^2$. **Implicit differentiation** concerns the differentiation of implicit functions. In this case, it actually makes our life easier if we turn $\ln(x)$ into an implicit function.

If I let $\ln(x) = y$, I can raise both sides to the power of e and get

$$\mathrm{e}^{\ln(x)} = \mathrm{e}^y. \tag{3.58}$$

Because logarithms are inverses of exponential,

$$x = \mathrm{e}^y. \tag{3.59}$$

By the rule of equality, (section 3.1), we can take the derivative of both sides w.r.t. y instead of x:

$$\begin{aligned}\frac{\mathrm{d}x}{\mathrm{d}y} &= \frac{\mathrm{d}\mathrm{e}^y}{\mathrm{d}y}\\ \frac{\mathrm{d}x}{\mathrm{d}y} &= \mathrm{e}^y.\end{aligned}$$

But we're looking for the derivative of y (which is just $\ln(x)$) w.r.t. x, not the derivative of x w.r.t. y. Here's where Leibniz's notation comes into clutch: we can swap the numerator with the denominator for both sides then, substitute back $y = \ln(x)$:

$$\begin{aligned}\frac{\mathrm{d}\ln(x)}{\mathrm{d}x} &= \frac{1}{\mathrm{e}^{\ln(x)}}\\ \therefore \frac{\mathrm{d}\ln(x)}{\mathrm{d}x} &= \frac{1}{x},\end{aligned}$$

and there: the derivative of the natural logarithm is the reciprocal.

With the power of natural logarithms, we can actually go back at the derivative of n^x and finally uncover the mystery behind the proportionality term that's lingering around. Start with the manipulation of n^x.

$$n^x = \left(\mathrm{e}^{\ln(n)}\right)^x = \mathrm{e}^{x\ln(x)}. \tag{3.60}$$

With the chain rule (section 3.1), let $u = x\ln(n)$:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}(n^x) &= \frac{\mathrm{d}\mathrm{e}^{x\ln(n)}}{\mathrm{d}x}\\ &= \frac{\mathrm{d}\mathrm{e}^u}{\mathrm{d}x} \times \frac{\mathrm{d}u}{\mathrm{d}u}\\ &= \frac{\mathrm{d}\mathrm{e}^u}{\mathrm{d}u} \times \frac{\mathrm{d}u}{\mathrm{d}x}\\ &= \mathrm{e}^{x\ln(n)} \times \frac{\mathrm{d}x\ln(n)}{\mathrm{d}x}\\ &= n^x\ln(n).\end{aligned}$$

And here it is: the mystery proportionality constant. It's just a consequence of the natural logarithm. Thus, one way to define the natural log would be

$$\ln(n) = \lim_{h\to 0}\frac{n^h - 1}{h}. \tag{3.61}$$

The antiderivative of other bases exponents are then given by

$$\int n^x = \frac{1}{\ln(n)}n^x + C. \tag{3.62}$$

3.5.1 The product rule and the quotient rule

Sometimes, we have to multiply the two functions together before taking the derivatives. There are two ways to do this. To keep the spirit of visualization, I shall first introduce the geometrical way, then the analytical way.

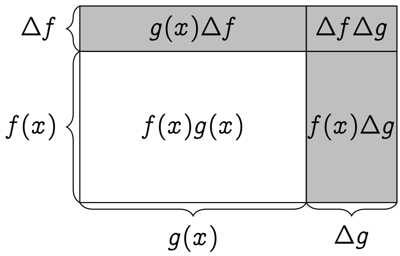


FIG. 3.4 | THE GEOMETRICAL INTERPRETATION OF THE PRODUCT RULE.

The derivative of $f(x)g(x)$ w.r.t. x can be thought of a rectangle with side length that's governed by $f(x)$ and $g(x)$. As shown in fig. 3.4, the area increase on side $f(x)$ is $g(x)\Delta f$ and on $g(x)$, $f(x)\Delta g$. The $\Delta f \Delta g$ part is basically negligible. Therefore,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\big(f(x)g(x)\big) &= \lim_{h\to 0}\frac{g(x)\Delta f + f(x)\Delta g}{h}\\
&= \lim_{h\to 0}\frac{f(x)(g(x+h)-g(x)) + g(x)(f(x+h)-f(x))}{h}\\
&= f(x)\lim_{h\to 0}\frac{g(x+h)-g(x)}{h} + g(x)\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}\\
&= f(x)\frac{\mathrm{d}g(x)}{\mathrm{d}x} + g(x)\frac{\mathrm{d}f(x)}{\mathrm{d}x}.
\end{aligned}$$

Which is what we call the product rule. Notice the alternation between f and g in the two terms. The derivative of $f(x)$ is multiplied by $g(x)$, and the derivative of $g(x)$ is multiplied by $f(x)$. This is a direct consequence of the diagram: the change in $f(x)$ is multiplied by $g(x)$ to give the area and also the other way around. You could check this with the method of increments, and it would still be true. I encourage you to do it.

To take derivatives of quotients of functions, just plug in $1\,/\,g(x)$

instead of $g(x)$. The final form should be

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{f(x)}{g(x)}\right) = \frac{f(x)\dfrac{\mathrm{d}g(x)}{\mathrm{d}x} - g(x)\dfrac{\mathrm{d}f(x)}{\mathrm{d}x}}{f(x)^2}. \tag{3.63}$$

But how's about the product of multiple functions? We can't really geometrically interpret those anymore. So we have to turn ourselves to the analytical method.

Let's think of this through. We don't know the derivative of products, but we know the derivative of *sums*: the linearity property. So if we can turn products into sum, this problem would be so easy! Gladly, the logarithms can do exactly that. To make our life easier, we shall use the natural logarithm. Because I want to save space, let's write $a(x)b(x)c(x)$ as just abc. Do know that these functions are all dependent on x. Start off with a manipulation of products.

$$\begin{aligned}\frac{\mathrm{d}(abc\ldots)}{\mathrm{d}x} &= \frac{\mathrm{d}}{\mathrm{d}x}(\mathrm{e}^{\ln(abc\ldots)}) \\ &= \frac{\mathrm{d}}{\mathrm{d}x}(\mathrm{e}^{\ln(a)+\ln(b)+\ln(c)+\ldots}).\end{aligned}$$

Now, let $\ln(a)+\ln(b)+\ln(c)+\cdots = u$ and use the chain rule,

$$\begin{aligned}&= \frac{\mathrm{d}}{\mathrm{d}x}(\mathrm{e}^{u})\cdot\frac{\mathrm{d}u}{\mathrm{d}u} \\ &= \frac{\mathrm{d}}{\mathrm{d}x}(\mathrm{e}^{u})\cdot\frac{\mathrm{d}}{\mathrm{d}x}(\ln(a)+\ln(b)+\ln(c)+\ldots) \\ &= \mathrm{e}^{u}\left(\frac{\mathrm{d}\ln(a)}{\mathrm{d}x}+\frac{\mathrm{d}\ln(b)}{\mathrm{d}x}+\frac{\mathrm{d}\ln(c)}{\mathrm{d}x}+\ldots\right).\end{aligned}$$

Then use the chain rule again on the terms in the parenthesis

$$\begin{aligned}&= \mathrm{e}^{u}\left(\frac{\mathrm{d}\ln(a)}{\mathrm{d}x}\frac{\mathrm{d}a}{\mathrm{d}a}+\frac{\mathrm{d}\ln(b)}{\mathrm{d}x}\frac{\mathrm{d}b}{\mathrm{d}b}+\frac{\mathrm{d}\ln(c)}{\mathrm{d}x}\frac{\mathrm{d}c}{\mathrm{d}c}+\ldots\right) \\ &= (abc\ldots)\left(\frac{\mathrm{d}\ln(a)}{\mathrm{d}a}\frac{\mathrm{d}a}{\mathrm{d}x}+\frac{\mathrm{d}\ln(b)}{\mathrm{d}b}\frac{\mathrm{d}b}{\mathrm{d}x}+\frac{\mathrm{d}\ln(c)}{\mathrm{d}c}\frac{\mathrm{d}c}{\mathrm{d}x}+\ldots\right) \\ &= (abc\ldots)\left(\frac{1}{a}\frac{\mathrm{d}a}{\mathrm{d}x}+\frac{1}{b}\frac{\mathrm{d}b}{\mathrm{d}x}+\frac{1}{c}\frac{\mathrm{d}c}{\mathrm{d}x}+\ldots\right).\end{aligned}$$

And there we have it: the generalized product rule.

3.5.2 Alternative derivations for the power rule

The power rule can also be derived using the same technique we just used. However, we use a different property of logarithm: $\ln(x^n) = n\ln(x)$.

$$\frac{\mathrm{d}}{\mathrm{d}x}x^n = \frac{\mathrm{d}}{\mathrm{d}x}\mathrm{e}^{n\ln(x)}.$$

Let $u = n\ln(x)$ then use the chain rule

$$\begin{aligned}
&= \frac{\mathrm{d}\mathrm{e}^u}{\mathrm{d}x} \cdot \frac{\mathrm{d}u}{\mathrm{d}u} \\
&= \frac{\mathrm{d}\mathrm{e}^u}{\mathrm{d}u} \cdot \frac{\mathrm{d}u}{\mathrm{d}x} \\
&= \mathrm{e}^u \cdot \frac{\mathrm{d}}{\mathrm{d}x}\big(n\ln(x)\big) \\
&= x^n \cdot n\frac{1}{x} = nx^{n-1}.
\end{aligned}$$

3.6 Implicit differentiation

Implicit differentiation also shows up in problems where two rate of changes are related to each other: the **related rates** problem. Consider a sliding ladder that's 5m long (fig. 3.5). At a certain time t, what's the sliding rate along the x-axis w.r.t. the y-axis?

The problem is asking for $\frac{\mathrm{d}x}{\mathrm{d}y}$. Here, the rate of sliding along the y-axis $\frac{\mathrm{d}y}{\mathrm{d}t}$ and along the x-axis $\frac{\mathrm{d}x}{\mathrm{d}t}$ is related because the ladder length is fixed. If y decreases, x must increase. The Pythagorean theorem says

$$x^2 + y^2 = 5^2. \tag{3.64}$$

We can take the derivative w.r.t. t on both side then use the chain rule

$$\begin{aligned}
\frac{\mathrm{d}x^2}{\mathrm{d}t} + \frac{\mathrm{d}y^2}{\mathrm{d}t} &= \frac{\mathrm{d}(5^2)}{\mathrm{d}t} \\
\frac{\mathrm{d}x^2}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}t} + \frac{\mathrm{d}y^2}{\mathrm{d}y}\frac{\mathrm{d}y}{\mathrm{d}t} &= 0
\end{aligned}$$

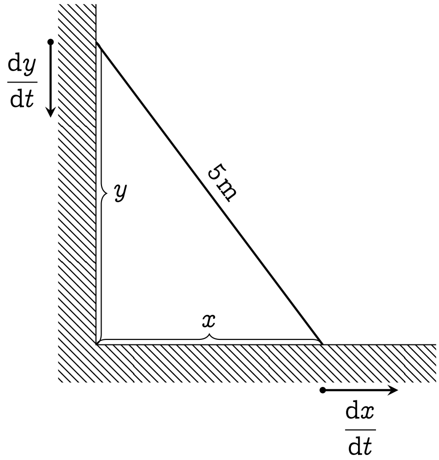


FIG. 3.5 | A LADDER LENGTH 5m SLIDING DOWN A CORNER.

$$2x\frac{\mathrm{d}x}{\mathrm{d}t} + 2y\frac{\mathrm{d}y}{\mathrm{d}t} = 0.$$

We can take advantage of the Leibniz's notation and multiply by $\mathrm{d}t$ on both sides giving

$$2x\,\mathrm{d}x + 2y\,\mathrm{d}y = 0.$$

To find $\frac{\mathrm{d}x}{\mathrm{d}y}$, we just have to isolate the variables,

$$\frac{\mathrm{d}x}{\mathrm{d}y} = -\frac{y}{x}.$$

And this should actually make sense. Because if y is increasing by a bit, x must decrease by some amount, and that amount is $-y\,/\,x$: the higher the y, the larger the rates of sliding.

3.7 Technique of integraton: substitution

Before we finish this chapter off, I'd want to take a look at a very useful technique for dealing with complicated integrals: substitution.

Substitution is a technique to simplify algebraic expressions. I believe you've already seen a multitude of them. Basically, we substitute some expression of x with u, then use everything in our power to change every x to u, including the differential element: from $\mathrm{d}x$ to $\mathrm{d}u$.

u-substitution works when the integrals are of the form

$$\int f(g(x)) \frac{\mathrm{d}g(x)}{\mathrm{d}x} \,\mathrm{d}x. \tag{3.65}$$

In this notation, it's easy to see that the $\mathrm{d}x$ will cancel, leaving us with

$$\int f(g(x)) \,\mathrm{d}g(x). \tag{3.66}$$

If $g(x) = u$, then our integral simplifies to

$$\int f(u) \,\mathrm{d}u. \tag{3.67}$$

Let's apply this knowledge to solve some integrals. Starting with the simplest: polynomic integrands.

Example 3.7.1: $\int x(x^2+1)^{10} \,\mathrm{d}x$

If $f(x) = x^{10}$, and $u = x^2 + 1$, then

$$\int x(x^2+1)^{10} \,\mathrm{d}x = \int u^{10} x \,\mathrm{d}x. \tag{3.68}$$

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

$$\int u^{10} x \,\mathrm{d}x = \int u^{10} \frac{1}{2}\frac{\mathrm{d}(x^2+1)}{\mathrm{d}x} \,\mathrm{d}x = \frac{1}{2}\int u^{10} \,\mathrm{d}(x^2+1) \tag{3.69}$$

Since $u = x^2 + 1$, our integral simplifies to

$$\frac{1}{2}\int u^{10} \,\mathrm{d}u. \tag{3.70}$$

Using the reversed power rule (eq. (3.20)), we get

$$\frac{1}{2}\frac{u^{11}}{11} = \frac{u^{11}}{22} = \frac{(x^2+1)^{11}}{22} + C. \tag{3.71}$$

Example 3.7.2: $\int x^2 \mathrm{e}^{4x^3+4}\,\mathrm{d}x$

The integrand is a composite function: $f(u) = \mathrm{e}^u$ and $u = 4x^3 + 4$. In which, $\frac{\mathrm{d}u}{\mathrm{d}x} = 12x^2$; therefore,

$$\int x^2 \mathrm{e}^{4x^3+4}\,\mathrm{d}x = \frac{1}{12}\int \mathrm{e}^u 12x^2\,\mathrm{d}x \tag{3.72}$$

$$= \frac{1}{8}\int \mathrm{e}^u \frac{\mathrm{d}u}{\mathrm{d}x}\,\mathrm{d}x \tag{3.73}$$

$$= \frac{1}{8}\int \mathrm{e}^u\,\mathrm{d}u = \frac{1}{8}\mathrm{e}^u \tag{3.74}$$

$$= \frac{1}{12}\mathrm{e}^{4x^3+4} + C. \tag{3.75}$$

Notice any patterns? If the integral has a polynomial u with degree n nested inside another function, there must be a term outside with the degree $n-1$ to substitute with. In **??** 3.7.1, $u = x^2 + 1$ and there's an x multiplied outside waiting to be substituted. In **??** 3.7.3, $u = 4x^3 + 4$, and there's an x^2 outside. This pattern is a consequence of the power rule, and is pretty common. So, here are some practice problems with answers and hints on the right.[3] Try not to look at it much.

1. $\int 3x^2\sqrt{x^3+5}\,\mathrm{d}x$ — $\frac{2}{3}(x^3+5)^{\frac{3}{2}} + C$, $u = x^3 + 5$
2. $\int \sqrt{4+3x}\,\mathrm{d}x$ — $\frac{2}{9}(3x+4)^{\frac{3}{2}} + C$, $u = 4 + 3x$
3. $\int \frac{24x}{(4x^2+4)^2}\,\mathrm{d}x$ — $\frac{3}{4x^2+4} + C$, $u = 4x^2 + 4$

Now let's move on from polynomials to other functions that are nested inside another function. For the substitution to work, there must be the function's derivative multiplied, waiting to be substituted. A classic example is:

[3] Practice problems taken from [2].

Example 3.7.3: $\displaystyle\int \frac{\ln(x)}{x}\,\mathrm{d}x$

This integral is not quite a composite function, but rather just a product between a function and its derivative. If $u = \ln(x)$, then $\dfrac{\mathrm{d}u}{\mathrm{d}x} = \dfrac{1}{x}$; therefore,

$$\int \frac{\ln(x)}{x}\,\mathrm{d}x = \int u \frac{\mathrm{d}u}{\mathrm{d}x}\,\mathrm{d}x \tag{3.76}$$

$$= \int u\,\mathrm{d}u \tag{3.77}$$

$$= \frac{u^2}{2} = \frac{\ln(x)^2}{2} + C. \tag{3.78}$$

Example 3.7.4: $\displaystyle\int x\mathrm{e}^{x^2}(\mathrm{e}^{x^2}+3)\,\mathrm{d}x$

Let $\mathrm{e}^{x^2} + 3 = u$, then

$$\frac{\mathrm{d}u}{\mathrm{d}x} = \frac{\mathrm{d}\mathrm{e}^{x^2}+3}{\mathrm{d}x} \tag{3.79}$$

$$= \frac{\mathrm{d}\mathrm{e}^{x^2}}{\mathrm{d}x} \tag{3.80}$$

Let $x^2 = v$, then by the chain rule,

$$= \frac{\mathrm{d}\mathrm{e}^{v}}{\mathrm{d}x} \cdot \frac{\mathrm{d}v}{\mathrm{d}v} \tag{3.81}$$

$$= \frac{\mathrm{d}\mathrm{e}^{v}}{\mathrm{d}v} \cdot \frac{\mathrm{d}x^2}{\mathrm{d}x} \tag{3.82}$$

$$= 2x\mathrm{e}^{v} = 2x\mathrm{e}^{x^2}. \tag{3.83}$$

Thus,

$$\int x\mathrm{e}^{x^2}(\mathrm{e}^{x^2}+3)\,\mathrm{d}x = \frac{1}{2}\int u\left(2x\mathrm{e}^{x^2}\right)\mathrm{d}x \tag{3.84}$$

$$= \frac{1}{2}\int u\frac{\mathrm{d}u}{\mathrm{d}x}\,\mathrm{d}x = \frac{1}{2}\int u\,\mathrm{d}u \tag{3.85}$$

$$= \frac{1}{2}\frac{u^2}{2} = \frac{(\mathrm{e}^{x^2}+3)^2}{4} + C. \tag{3.86}$$

Sometimes, they need some *creative* algebraic manipulations, and sometimes even multiple substitution. E.g.,[4]

Example 3.7.5: $\frac{1}{x(x^5+1)}\,\mathrm{d}x$

If we want to use power rule substitution, there's only x^5 but no x^4 to substitute with. So, we need to get a bit creative.

First, factor the x^5 out of the inner parenthesis and get

$$\int \frac{1}{x^6\left(1+\frac{1}{x^5}\right)}\,\mathrm{d}x. \tag{3.87}$$

Now, we can let $u = 1+\frac{1}{x^5}$. By the power rule, $\frac{\mathrm{d}u}{\mathrm{d}x} = -\frac{5}{x^6}$. Our integral then becomes

$$\begin{aligned} -5\int \frac{1}{u}\left(-\frac{5}{x^6}\right)\mathrm{d}x &= -\frac{1}{5}\int \frac{1}{u}\frac{\mathrm{d}u}{\mathrm{d}x}\,\mathrm{d}x && (3.88)\\ &= -\frac{1}{5}\int \frac{1}{u}\,\mathrm{d}u = -5\ln(x) && (3.89)\\ &= -\frac{1}{5}\ln\left(1+\frac{1}{x^5}\right)+C && (3.90) \end{aligned}$$

In certain problems, there aren't any outside terms to substitute with, so you have to create them. E.g.,

Example 3.7.6: $\int \left(1+\sqrt{x}\right)^4\mathrm{d}x$

We want to get rid of the term inside the parenthesis. So, let $u = 1+\sqrt{x}$, and by the power rule, $\frac{\mathrm{d}u}{\mathrm{d}x} = \frac{1}{2\sqrt{x}}$. However, there is no $\frac{1}{\sqrt{x}}$ outside to substitute with. Let's be adventurous and substitute u in

[4] Taken from MIT Integration Bee 2019, I_9 [1]

anyways and see what happens.

$$\int (1+\sqrt{x})^4 \, \mathrm{d}x = \int u^4 \, \mathrm{d}x \tag{3.91}$$

If $\dfrac{\mathrm{d}u}{\mathrm{d}x} = \dfrac{1}{2\sqrt{x}}$, then $\mathrm{d}x = 2\sqrt{x}\,\mathrm{d}u$,

$$= \int u^4 \left(2\sqrt{x}\right) \mathrm{d}u. \tag{3.92}$$

We're almost there, now we just have to replace the leftover x with u. Since $u = 1+\sqrt{x}$, $\sqrt{x} = u - 1$; thus,

$$\int u^4 \left(2\sqrt{x}\right) \mathrm{d}u = 2\int u^4 (u-1)\,\mathrm{d}u \tag{3.93}$$

$$= 2\int u^5 - u^4 \,\mathrm{d}u = 2\left(\frac{u^6}{6} - \frac{u^5}{5}\right) \tag{3.94}$$

$$= 2\left(\frac{(1+\sqrt{x})^6}{6} - \frac{(1+\sqrt{x})^5}{5}\right). \tag{3.95}$$

These problems definitely take a while to get used to. So I highly encourage you to practice integrating using a problem set ([1], [2]). The unfortunate thing is that most problem sets include trigonometry. So, you'd have to pick the one without them to practice with. Or, you can just wait until we discuss about trigonometry and then practice from there.

3.8 Integral of the reciprocal

By the fundamental theorem of calculus (theorem 1), the antiderivative of $\dfrac{1}{x}$ is $\ln(x)$. You might believe this, and it's not wrong. But I find it very disturbing and unresolved. It's like taking the result and pointing it back to the origin. It doesn't really make sense. So from this dissatisfaction, I spent a night coming up with a way to derive this using just the reversed power rule. Enjoy the transformation!

$$\begin{aligned}
\int \frac{1}{x}\,\mathrm{d}x &= \int \lim_{h\to 0}\left(\frac{1}{2}x^{-1+h} + \frac{1}{2}x^{-1-h}\right)\mathrm{d}x \\
&= \lim_{h\to 0}\int \left(\frac{1}{2}x^{-1+h} + \frac{1}{2}x^{-1-h}\right)\mathrm{d}x \\
&= \lim_{h\to 0}\int \left(\frac{1}{2}\frac{x^{-1+h+1}}{(-1+h+1)} + \frac{1}{2}\frac{x^{-1-h+1}}{(-1-h+1)}\right)\mathrm{d}x \\
&= \lim_{h\to 0}\left(\frac{1}{2}\frac{x^h}{h} - \frac{1}{2}\frac{x^{-h}}{-h}\right) = \lim_{h\to 0}\left(\frac{1}{2}\frac{x^h}{h}\frac{x^h}{h} - \frac{1}{2}\frac{1}{hx^h}\right) \\
&= \lim_{h\to 0}\left(\frac{x^{2h}-1}{2hx^h}\right) = \lim_{h\to 0}\left(\frac{\mathrm{e}^{2h\ln(x)}-1}{2h\ln(x)}\cdot\frac{\ln(x)}{x^h}\right) \\
&= \lim_{h\to 0}\left(\frac{\mathrm{e}^{2h\ln(x)}}{2h\ln(x)}\right)\cdot\lim_{h\to 0}\left(\frac{\ln(x)}{x^h}\right) \\
&= \lim_{h\to 0}\left(\frac{\mathrm{e}^{2h\ln(x)}}{2h\ln(x)}\right)\cdot\ln(x). \qquad (3.96)
\end{aligned}$$

Then, we evaluate the limit at the front by letting $u = 2h\ln(x)$. When $h \to 0$, $u \to 0$ as well. Then, use the definition of e from eq. (3.37).

$$\begin{aligned}
\lim_{h\to 0}\left(\frac{\mathrm{e}^{2h\ln(x)}}{2h\ln(x)}\right) &= \lim_{u\to 0}\left(\frac{\mathrm{e}^u-1}{u}\right) \\
&= \lim_{u\to 0}\left(\frac{\left(\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n\right)^u}{u}\right).
\end{aligned}$$

Change the limits from $n \to \infty$ into $n \to 0$. Notice, $\lim_{n\to 0}\left(1+\frac{1}{n}\right)^n = \lim_{n\to 0}(1+n)^{1/n}$. If $n \to 0$ and $u \to 0$, that means $n = u$. Substitute $n = u$ into the limit,

$$\begin{aligned}
&= \lim_{u\to 0}\left(\frac{\left(\lim_{u\to 0}(1+u)^{1/u}\right)^u - 1}{u}\right) \\
&= \lim_{u\to 0}\left(\frac{1+u-1}{u}\right) = 1.
\end{aligned}$$

Then, substitute this limit back into eq. (3.96), you'll see that

$$\int \frac{1}{x}\,\mathrm{d}x = \lim_{h\to 0}\left(\frac{\mathrm{e}^{2h\ln(x)}}{2h\ln(x)}\right)\cdot\ln(x) = \ln(x) + C,$$

completing the proof.

3.9 Formula for Chapter 3

3.9.1 Formula for derivatives of functions

1. $f(x) = g(x) \implies \dfrac{\mathrm{d}f(x)}{\mathrm{d}x} = \dfrac{\mathrm{d}g(x)}{\mathrm{d}x}$ (Rules of Equality)
2. For $c \in \mathbb{R}$, $\dfrac{\mathrm{d}(c)}{\mathrm{d}x} = 0$ (Derivative of a constant)
3. $\dfrac{\mathrm{d}x}{\mathrm{d}y} = \dfrac{\mathrm{d}x}{\mathrm{d}u} \times \dfrac{\mathrm{d}u}{\mathrm{d}y}$ (Chain rule)
4. $\dfrac{\mathrm{d}}{\mathrm{d}x}\big(af(x) + bg(x)\big) = a\dfrac{\mathrm{d}f(x)}{\mathrm{d}x} + b\dfrac{\mathrm{d}g(x)}{\mathrm{d}x}$ (Linearity of differentiation)
5. $\dfrac{\mathrm{d}}{\mathrm{d}x}(ax^n) = anx^{n-1}$ (Power rule)
6. $\dfrac{\mathrm{d}}{\mathrm{d}x}(n^x) = n^x \ln(n)$, $\dfrac{\mathrm{d}(\mathrm{e}^x)}{\mathrm{d}x} = \mathrm{e}^x$ (Derivative of exponentials)
7. $\dfrac{\mathrm{d}}{\mathrm{d}x}\big(\ln(x)\big) = \dfrac{1}{x}$ (Derivative of natural logarithms)
8. $\dfrac{\mathrm{d}}{\mathrm{d}x}\big(f_1(x)f_2(x)\big) = f_1(x)\dfrac{\mathrm{d}f_2}{\mathrm{d}x} + f_2(x)\dfrac{\mathrm{d}f_1}{\mathrm{d}x}$ (Product rule for two functions)
9. $\dfrac{\mathrm{d}}{\mathrm{d}x}(f_1 f_2 \ldots f_n) = f_1 f_2 \ldots f_n\left(\dfrac{1}{f_1}\dfrac{\mathrm{d}f_1}{\mathrm{d}x} + \dfrac{1}{f_2}\dfrac{\mathrm{d}f_2}{\mathrm{d}x} + \cdots + \dfrac{1}{f_n}\dfrac{\mathrm{d}f_n}{\mathrm{d}x}\right)$ (Generalized product rule)

3.9.2 Formula for antiderivatives of functions

1. $f(x) = g(x) \implies \displaystyle\int f(x)\,\mathrm{d}x = \int g(x)\,\mathrm{d}x$ (Rules of Equality)
2. $\displaystyle\int af(x) + bg(x)\,\mathrm{d}x = a\int f(x)\,\mathrm{d}x + b\int g(x)\,\mathrm{d}x$ (Linearity of integration)
3. $n \neq -1$, $\displaystyle\int ax^n\,\mathrm{d}x = a\frac{x^{n+1}}{n+1} + C$ (Reversed power rule)

4. $\displaystyle\int n^x \, \mathrm{d}x = \frac{1}{\ln(n)} n^x + C$, $\displaystyle\int \mathrm{e}^x \, \mathrm{d}x = \mathrm{e}^x + C$ (Antiderative of exponentials)

5. $\displaystyle\int \frac{1}{x} \, \mathrm{d}x = \ln(x) + C$ (Antiderivative of the reciprocal)

6. $\displaystyle\int f(g(x)) \frac{\mathrm{d}g}{\mathrm{d}x} \, \mathrm{d}x = \int f(u) \, \mathrm{d}u; u = g(x)$ (u-Substitution)

3.9.3 Definition for various functions and constants

1. $\displaystyle \mathrm{e}^x = \lim_{k \to 0} \sum_{i \to 0}^{k} \frac{1}{i!} = 1 + \frac{x^1}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots = \lim_{n \to \infty} \left(1 + \frac{x}{n}\right)^n$

2. $\displaystyle \mathrm{e} = \sum_{i=0}^{\infty} \frac{1}{i!} = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \cdots = \lim_{n \to \infty} \left(1 + \frac{1}{n}\right)^n$

3. $\displaystyle \ln(x) = \lim_{n \to 0} \left(\frac{x^h - 1}{h}\right)$

CHAPTER 4

A study of graphs with calculus

4.1 Invitation: sketching mountain ranges

If I told you to *sketch* out x^2, it'd be easy, right? Just a simple parabola would do. But if I say: plot $x^4 - 2x^3 + x^2 - x + 1$? It isn't so easy now is it. You could substitute various x and plot it into the graph directly. But I'm asking for a rough sketch, not a plot. I just want the general shape and structure of the function: where are the highest and lowest points, the x-intercept, and the y-intercept. We're going to learn how to do exactly that in this chapter.

Imagine a function as a vast mountain range. When sketching this landscape, we're interested in identifying its peaks, valleys, and plateaus. A peak represents the sharp summit of a mountain, where the land falls away in all directions; a valley is a sunken area, nestled lower than the surrounding terrain; and a plateau is a level stretch, a temporary calm before the ascent resumes toward higher peaks or descends into deeper valleys. Standing at a peak, the land around you dips downward. In contrast, in a valley, everything rises above you. On a plateau, the ground is flat under your feet, but when gazed into the distance, one side raises higher, but the other reveals the path down. These important features are all characterized by a flat ground: a place with zero slope. If the mountain range represents function, then all of these points have zero first derivative.

Still, the information about zero first derivative is not sufficient to determine whether the point is a peak, a valley, or a plateau. What helps classify these points is the sign of the second derivative.

The second derivative suggest the concavity of a graph at that point. If the first derivative gives you the slope, the second derivative will give you the rate of change of that slope. If the second derivative is positive, it means that the function is increasing around that point; thus, that point must be a valley. If it's negative, then everything is decreasing around; thus, it's a peak. If it's zero, it's simply a plateau.

Just like a mountain, a function can have many points that has zero derivatives. Functions with more than one peaks exists. The first derivative does not provide us with any information on which one of those peaks are the highest. It only tells you where the peak is. In order to find which one is the highest, we'd just have to directly find the function's value at those points, then compare it.

In mathematics terminology, the points that have zero first derivative is called a **stationary point**: the value of the function around it is stationary. The valleys are called, **local minima**; the peaks, **local maxima**, and the plateau, **inflection point**. The word *local* suggests that locally, these are either the maximum or minimum point, but it might not be globally. Like how you can stand on your house and say that "this is the highest point of my house," but you're still not higher than mount everest. We call the global maximum or minimum point, the **absolute maximum** and **absolute minimum**.

4.2 Minima and maxima

This section focuses on finding the minima and maxima of functions and sketching it. These methods are applicable for all kinds of function, but we'll only study the important ones: quadratics and cubics.

4.2.1 Shape of quadratics

Let's start off with quadratics: a function with only one peak or valley. For any given quadratic

$$f(x) = ax^2 + bx + c, \tag{4.1}$$

Its derivative can be found using the power rule (eq. (3.20)).

$$\frac{\mathrm{d}}{\mathrm{d}x} f(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left(ax^2 + bx + c\right) \tag{4.2}$$

$$= 2ax + b. \tag{4.3}$$

The peak or valley of $f(x)$ is at $x = x_0$ where $\frac{\mathrm{d}}{\mathrm{d}x} f(x)\Big|_{x=x_0} = 0$. The notation $\frac{\mathrm{d}}{\mathrm{d}x} f(x)\Big|_{x=x_0}$ tells you that this derivative needs to be evaluated at $x = x_0$. For a quadratic, this means

$$2ax_0 + b = 0 \tag{4.4}$$

$$x_0 = -\frac{b}{2a}, \tag{4.5}$$

which matches with the usual result from traditional algebra. The y value of these points can be found by plugging in $x = -\frac{b}{2a}$ into $f(x)$.

$$f\left(-\frac{b}{2a}\right) = a\left(-\frac{b}{2a}\right)^2 + b\left(-\frac{b}{2a}\right) + c \tag{4.6}$$

$$= \frac{b^2}{4a} - \frac{b^2}{2a} + c \tag{4.7}$$
$$= c - \frac{3b^2}{4a} \tag{4.8}$$

Other important points include the y and x intercept. The y-intercept can easily be found by setting $x = 0$:

$$y_0 = a(0)^2 + b(0) + cy_0 = c. \tag{4.9}$$

Sadly, calculus does not offer any tools to find the exact root of an equation; you still have to use the quadratic equation

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \tag{4.10}$$

Classically, the direction that the quadratic The second derivative of a quadratic is determined by the sign of a. The same result can be derived from the second derivative.

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2} f(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right) \tag{4.11}$$
$$= \frac{\mathrm{d}}{\mathrm{d}x}\big(2ax + b\big) \tag{4.12}$$
$$= 2a. \tag{4.13}$$

Since we only care about the sign, the two does not matter. This results tell us that if a is positive, it's a right side up parabola, and the stationary point is a local minima. If it isn't, then it's upsidedown and the stationary point is a local maxima.

A parabola doesn't have any inflection point because it only occurs when the second derivative is zero. However, the only way that $\frac{\mathrm{d}^2}{\mathrm{d}x^2} f(x) = 0$ is for a to be zero. If it happens, then $f(x)$ isn't a parabola anymore, but rather a straight line.

Example 4.2.1: Sketch the graph of $f(x) = -2x^2 + x + 6$

Because a is negative, the parabola is upsidedown. The vertex is at

$$x = -\frac{b}{2a} = -\frac{1}{2(-2)} = \frac{1}{4}, \tag{4.14}$$

$$y = -2\left(\frac{1}{4}\right)^2 + \left(\frac{1}{4}\right) + 6 = 6.125. \tag{4.15}$$

The y-intercept is trivially 6. The x intercept are the roots of this polynomial, given by

$$x = \frac{-(1) \pm \sqrt{(1)^2 - 4(-2)(6)}}{2(-2)} = 2, -\frac{3}{2}. \tag{4.16}$$

Factorization would also yield the same result.

$$-2x^2 + x + 6 = -(2x^2 - x - 6) \tag{4.17}$$

$$= -(x-2)(2x+3) \to x = 2, -\frac{3}{2} \tag{4.18}$$

Therefore, the sketched graph would be

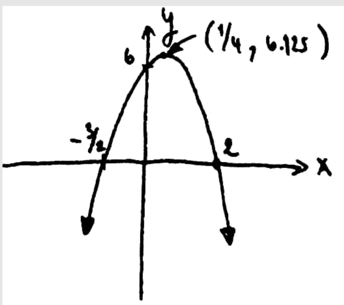


FIG. 4.1 | SKETCH OF THE FUNCTION $-2x^2 + x + 6$.

We can also determine the amount of roots by just looking at the y-value of the vertex point, given by eq. (4.8), and the direction of the parabola. But it's more computationally expensive than the determinant method which says,

> A parabola has no roots if $b^2 - 4ac < 0$, one root if $b^2 - 4ac = 0$, and two roots if $b^2 - 4ac > 0$.

This is a direct result of the quadratic formula: $x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$ only have real solution if the terms in the root, $b^2 - 4ac$, is positive.

4.2.2 Shape of cubics

Cubics are third order polynomials

$$f(x) = ax^3 + bx^2 + cx + d. \tag{4.19}$$

It's one of the many family of function that can contain more than one stationary point. The x coordinate of the stationary points can be found by using the same method as earlier.

$$\frac{\mathrm{d}f(x)}{\mathrm{d}x} = \frac{\mathrm{d}}{\mathrm{d}x} ax^3 + bx^2 + cx + d \tag{4.20}$$

$$= 3ax^2 + 2bx + c, \tag{4.21}$$

in which

$$\left.\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right|_{x=x_0} = 0 = 3ax_0^2 + 2bx_0 + c \tag{4.22}$$

$$x_0 = \frac{-2b \pm \sqrt{(2b)^2 - 4(3a)(c)}}{2(3a)} \tag{4.23}$$

$$= \frac{-2b \pm 2\sqrt{b^2 - 3ac}}{6a} \tag{4.24}$$

$$= \frac{-b \pm \sqrt{b^2 - 3ac}}{3a}. \tag{4.25}$$

By the terms under the root, a cubic has no stationary point if $\frac{b^2 - 3ac}{3a} < 0$, one if $\frac{b^2 - 3ac}{3a} = 0$, and two if $\frac{b^2 - 3ac}{3a} > 0$.

The function $f(x)$ intersects the y axis when $x = 0$:

$$f(0) = a(0)^3 + b(0)^2 + c(0) + d = d; \tag{4.26}$$

thus, the y intercept is at $(0, d)$. For the x-intercept, again, there is no tool in calculus that can be used to find the exact root. So, you'd just have to either factor the function, or use the very lengthy cubic formula.

$$x = \sqrt[3]{q + \sqrt{q^2 + (r - p^2)^3}} + \sqrt[3]{q - \sqrt{q^2 + (r - p^2)^3}} + p \tag{4.27}$$

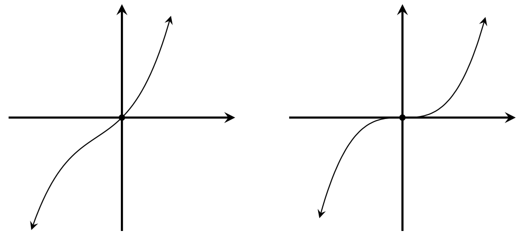
where

$$p = -\frac{b}{3a}, \quad q = p^3 + \frac{bc - 3ad}{6a^2} \quad \text{and,} \quad r = \frac{c}{3a}. \tag{4.28}$$

Another topic that I'd like to discuss is the amount of root of cubics. Normally, we would use factorization to determine the amount. From the shape of the cubic, one side shoots of to positive infinity, and the other, negative infinity. And by continuity, it must intersect the x axis somewhere inbetween and create a root. Therefore, all cubics must have at least one root over the reals. But what if we want to know whether it has a second or third roots or not? We can just look at the amount and the position of the stationary point, given by eq. (4.25).

A key feature of stationary points is that they indicate where a function changes direction. If a function is increasing and then begins to decrease, there must be a stationary point marking this transition. So, if a function lacks a stationary point, and one end tends toward positive infinity while the other falls to negative infinity, we can conclude the function is always either increasing or decreasing. Once the function intersects the x-axis to form a root, it cannot turn back to create another root.

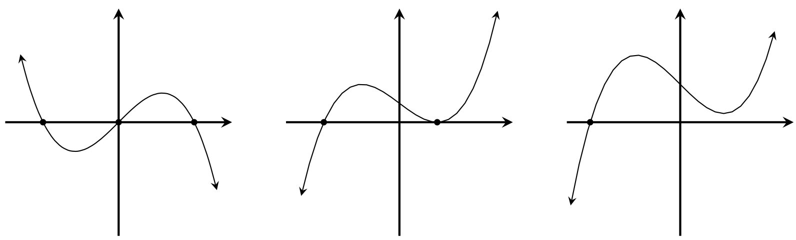
A cubic with zero stationary points ($\frac{b^2-3ac}{3a} < 0$) are such kind of function. One side goes to infinity, and the other, negative infinity. Since it has no stationary point, it must only have one root. One such example is the function $f(x) = x^3 + x^2 + x$, plotted in fig. 4.2a.



(A) x^3+x^2+x: No vertex (B) x^3: One vertex

Fig. 4.2 | Cubics with zero and one vertex

Similarly, cubics with only one stationary point ($\frac{b^2-3ac}{3a}=0$) only have one root. Illustrated in fig. 4.2b, the stationary point must be an inflection point, as it cannot represent a local maximum or minimum. This is due to the behavior of the function, which extends toward both positive and negative infinity. If the stationary point were a local maximum or minimum, it would turn into a parabola, and thus needs another point to reverse its direction again to be a cubic. Hence, cubic functions with only one stationary point have just one root.



(A) x^3+x: Three roots (B) $x^3-\frac{3}{4}x+\frac{1}{4}$: Two roots (C) x^3-x: One root

Fig. 4.3 | Two vertices cubics with varying amount of roots

For cubics with two vertex, one of those must be a local maximum, the other, a local minimum. The amount of roots can be directly determined by the position of these points. A cubic will have three roots if the vertex aren't on the same side of the x axis, it must have three roots, illustrated in fig. 4.3a. If one of the vertex is at $y=0$, then it has two roots;

one of them being the vertex itself, illustrated in fig. 4.3b. If both vertices are on the same side, then it must have only one root; illustrated in fig. 4.3c. It is left as an exercise for the reader to find the position of vertices of the cubics given in fig. 4.3

4.2.3 Optimization problems

The method of finding local maxima, minima, and inflection points are very useful for sketching graphs. Are there any real life applications to this? The answer is in optimization.

Optimization is about finding the best solution for a problem. What's the best time to harvest the crop? What's the best placement for an air conditioner? What's the best way to make profit?, etc. These problems are all about finding the maximum values of some variables. That's exactly what our tools have been able to do: finding the maxima and minima of some functions. Thus, I'd like to discuss about some applications of maxima and minima on some physical problems.

Classic maximum area problem. Given a rope length L, what's the maximum rectangular area that this rope can enclose?

To solve this problem, we start by constructing our rectangle. Drawn in fig. 4.4, the perimeter of a rectangle is given by $2(a+b) = L$. Without loss of generality, let side a has length l. Thus,

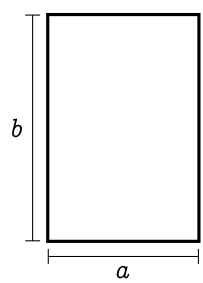


FIG. 4.4 | A RECTANGLE FOR OPTIMIZATION

$$2(l+b) = L \tag{4.29}$$

$$b = \frac{L}{2} - l \tag{4.30}$$

The area of the rectangle, which is a function of l, $A(l)$ is then

$$A(l) = a \times b = l\left(\frac{L}{2} - l\right) = -l^2 + l\frac{L}{2}. \tag{4.31}$$

Since this function is a parabola, its stationary point must be the local maxima, and must also be the global maxima. This point is where our function is the highest. It can be found by using the derivative w.r.t. l.

$$\frac{\mathrm{d}}{\mathrm{d}l}A(l) = -2l + \frac{L}{2} \tag{4.32}$$

Using the same argument as earlier, we want to find the point l_0 where $\frac{\mathrm{d}}{\mathrm{d}l}A(l)\Big|_{l=l_0} = 0$:

$$0 = -2l_0 + \frac{L}{2} \tag{4.33}$$

$$l_0 = \frac{L}{4}; \tag{4.34}$$

thus, the maximum $A(l)$ is

$$A\left(\frac{L}{4}\right) = -\left(\frac{L}{4}\right)^2 + \left(\frac{L}{4}\right)\frac{L}{2} = \frac{L^2}{16} \tag{4.35}$$

Funnily enough, the maximum area that a perimeter length L can increase is also a parabola $\frac{1}{16}L^2$.

Modified maximum area problem. Given a fence length L. What's the maximum area that I can enclose, given that one of the sides doesn't need any fence?

If our rectangle has sides a and b, and one side doesn't need a fence, without loss of generality, the perimeter is $a + 2b = L$. If $a = l$, then $b = \frac{L-l}{2}$. The area of the rectangle then becomes

$$A(l) = a \times b = l\left(\frac{L-l}{2}\right) = -\frac{1}{2}l^2 + \frac{L}{2}l. \tag{4.36}$$

Using the same argument,

$$\frac{\mathrm{d}}{\mathrm{d}l}A(l) = -l + \frac{L}{2} \tag{4.37}$$

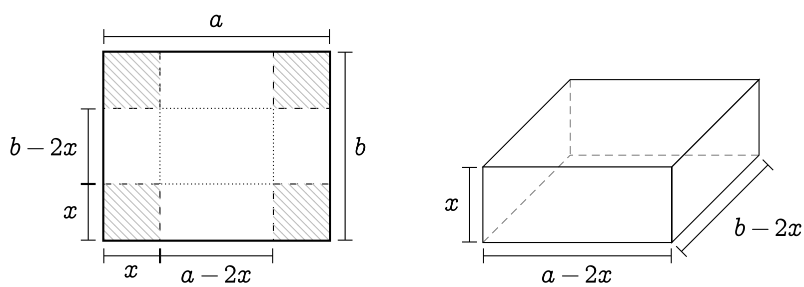
$$\frac{\mathrm{d}}{\mathrm{d}l}A(l)\bigg|_{l=l_0} = 0 = -l_0 + \frac{L}{2} \tag{4.38}$$

$$l_0 = \frac{L}{2}, \tag{4.39}$$

and

$$A(l_0) = -\frac{1}{2}\left(\frac{L}{2}\right)^2 + \frac{L}{2}\left(\frac{L}{2}\right) = \frac{3}{8}L^2. \tag{4.40}$$

Volume of a box. Let's say you have a paper size $a \times b$, shown in fig. 4.5a. If you want to fold it up into a box without a closing top by cutting a square size x on each corner, shown in fig. 4.5b, find the maximum volume that the box can enclose.



(A) UNFOLDED. THE GRAY PART IS CUT OUT. **(B)** FOLDED

FIG. 4.5 | BOX VOLUME OPTIMIZATION

Shown in fig. 4.5a, if we cut out a square sidelength x from all corners, then the sidelength of the final box will be $a-2x$, $b-2x$, and x. The volume $V(x)$, is then

$$V(x) = x(a-2x)(b-2x) = 4x^3 + (-2a-2b)x^2 + (ab)x \tag{4.41}$$

This function is a cubic, and we can use eq. (4.25) to find the x coordinate of the vertex.

$$x_0 = \frac{-(-2a-2b) \pm \sqrt{(-2a-2b)^2 - 3(4)(ab)}}{3(4)} \tag{4.42}$$
$$= \frac{2a+2b \pm \sqrt{4a^2+4b^2-4ab}}{12} \tag{4.43}$$
$$= \frac{a+b \pm \sqrt{a^2+b^2-ab}}{6} \tag{4.44}$$

If $a^2+b^2-ab=0$, there's only one stationary point at $x_0 = \frac{a+b}{6}$, which is an inflection point. Otherwise, it must have two stationary points. Since the leading coefficient of $V(x)$ is positive, $V(x)$ must approach negative infinity when x approaches infinity. Meaning that, the highest local maxima is on the right vertex; thus, we select the negative root:

$$x_0 = \frac{a+b-\sqrt{a^2+b^2-ab}}{6}. \tag{4.45}$$

The full solution is left out due to arithmetical tediousness. It'd be better if you just plug in the numbers.

4.3 Tangent to a curve

In this section, I'll show you how to find a line that's tangent to a curve. But to do it the traditional way would be a bit boring, as it's quite dry and actually doesn't help us with graph sketching at all. So, I shall introduce it via the Newton-Raphson's root finding algorithm; as it's an algorithm that directly takes advantage of the line tangent to a function, and it shows up in many places. So much so that it exists inside the core of the fast square root algorithm which is implemented in every computer.

4.3.1 Newton-Raphson method

Polynomials appears everywhere. Most of the time, you're required to find its root. But it might come as a surprise for you that it is *impossible* to find a general for finding the root of polynomials with degrees higher than four. This is a direct consequence of Galois theory[1], one of the most important theory in abstract algebra. So most of the time, we resort to root approximation methods.

The root approximation method which is relevant to calculus is the Newton-Raphson's root finding algorithm. I'd try to explain this algorithm has much as I can, but I think it's better if you see it.

The Newton-Raphson algorithm takes advantage of slopes descent. Let's say you want to approximate the root of $f(x)$. You put down a random guess, say x_1. If that point is a root, $f(x_1) = 0$. If its not, the root is either above ($f(x_1) < 0$) the point or under ($f(x_1) > 0$). What you do next is that you draw a line tangent to the function $f(x)$ at $x = x_1$ and let it cross the x-axis, shown in fig. 4.6. Intuitively, the tangent line will intersect the x-axis at x_2 where x_2 should hopefully be closer to the root. Our new guess will then be $(x_2, f(x_2))$. We then repeat this method until the approximation are satisfactory.

Here, I shall use the Newton-Raphson root finding algorithm as a way to introduce you about how calculus can be further applied to geometry.

The first part of the Newton-Raphson method is to be able to find

[1]The Galois theory is a profound result from abstract algebra. You would usually learn in graduate mathematics, but if you want to just scratch the surface, I found a video by *Math Visualized*, on *Galois Theory Explained Simply* [3], that explains it quite clearly. Still, I need to watch it a couple of times before understanding it.

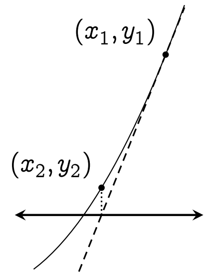


FIG. 4.6 | NEWTON-RAPHSON'S METHOD ILLUSTRATED

the tangent to a curve. Let's first start with a line equation

$$y = mx + c. \tag{4.46}$$

If my first guess is on the point $(x_1, f(x_1))$, the line that passes through that point and tangent to $f(x)$ is

$$y = \left.\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right|_{x=x_1} x + c. \tag{4.47}$$

where $m = \left.\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right|_{x=x_1}$. And c is the intersection with the x axis.

But we don't have to actually find c. For a line tangent to a curve, we just want to move the line from the origin to the tangent point $(x_1, f(x_1))$. To move the line to the right by x_1 units, we substitute x with $x - x_1$, and to shift it up by y_1, substitute y with $y - y_1$. Thus, the line equation turns into

$$y - y_1 = (x - x_1) \left.\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right|_{x=x_1} \tag{4.48}$$

$$y - f(x_1) = (x - x_1) \left.\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right|_{x=x_1}. \tag{4.49}$$

For simplicity of notation, I shall let $\mathrm{D}_x f(x_1)$ represent $\frac{\mathrm{d}f(x)}{\mathrm{d}x}$:

$$y - f(x_1) = (x - x_1)\,\mathrm{D}_x f(x_1). \tag{4.50}$$

Then, Newton-Raphson method tells us that the intersection between the tangent line and the x axis is what our next guess, x_2, should be. It can be easily evaluated by setting $y = 0$.

$$0 - f(x_1) = (x_2 - x_1)\,\mathrm{D}_x f(x_1) \tag{4.51}$$

$$x_2 = x_1 - \frac{f(x_1)}{\mathrm{D}_x f(x_1)} \tag{4.52}$$

Generally, we can repeat this forever. Thus,

$$x_{n+1} = x_n - \frac{f(x_n)}{\mathrm{D}_x f(x_n)} \tag{4.53}$$

a recurrence relation for finding roots of functions. Notice that we don't even need to know the y value for this algorithm to work. And that's the beauty of it.

Let's use this root finding method to approximate $\sqrt{2}$, which is the root of the polynomial $x^2 - 2 = 0$. Since the derivative of $x^2 - 2$ is $2x$, Newton-Raphson method (eq. (4.53)) gives

$$x_{n+1} = x_n - \frac{x_n^2 - 2}{2x_n} \tag{4.54}$$

$$= \frac{x_n^2 + 2}{2x_n}. \tag{4.55}$$

Since $\sqrt{2} \approx 1.414213\ldots$, let's use $x_1 = 1$ as our first guess.

$$x_2 = \frac{x_1^2 + 2}{2x_1} \tag{4.56}$$

$$= \frac{1^2 + 2}{2(1)} = 1.5, \tag{4.57}$$

which is closer to 1.41 than our first guess. Next,

$$x_3 = \frac{x_2^2 + 2}{2x_2} \tag{4.58}$$

$$= \frac{1.5^2 + 2}{2(1.5)} = 1.4166\ldots, \tag{4.59}$$

which is off by just a 0.2% error. By the forth iteration, we get $1.414215\ldots$, which is off by a mere 0.0002%. And we can stop here.

4.3.2 Numerical version of the Newton-Raphson method

For a function that you can't find a derivative, or you're just too lazy to explicitly find the derivative, you can turn the Newton-Raphson method into a numerical method via the definition of derivativeeq. (4.53).

$$x_{n+1} = x_n - \frac{f(x_n)}{\lim_{h\to 0}\frac{f(x_n+h)-f(x_n)}{h}} \tag{4.60}$$

$$= x_n - \lim_{h\to 0}\frac{f(x_n)h}{f(x_n+h)-f(x_n)}. \tag{4.61}$$

Computationally, we can set h to any arbitrarily small values: 0.01 for example. If you have some experiences with coding, this can be easily implemented by using a `for` loop. Here's some python code that I've written to do just that. Just put in the function in `def f(x):` block, put the amount of iteration in `iterationCount`, your first guess in `firstGuess`, and the value of h in `derivativeStep`. If you have some experiences, I highly encourage you to tinker around.

```python
import matplotlib.pyplot as plt
import math

def f(x):  # Insert function here
    return x**3 - 2 * x + 1

iterationCount = 10  # Amount of iteration
firstGuess = -2  # First guess
derivativeStep = 0.01  # For calculating the method of
↪ increments
approximationList = [float(firstGuess)]  # List of
↪ approximations
```

```python
for i in range(iterationCount):
    diff = (
        f(approximationList[i] + derivativeStep) -
         ↪ f(approximationList[i])
    ) / derivativeStep  # Finding the derivative at a
     ↪ certain point
    nextGuess = (
        approximationList[i] - f(approximationList[i]) /
         ↪ diff
    )  # Newton-Raphson's method
    approximationList.append(nextGuess)  # Add next guess to
     ↪ the list of approximations

print(approximationList)  # Print out the list of
 ↪ approximations
fig, axes = plt.subplots()  # Initiate plot
axes.plot(list(range(iterationCount + 1)),
 ↪ approximationList)  # Plotting
axes.set_xlabel("Iteration")  # Setting the x-axis label
axes.set_ylabel("Guess")  # Setting the y-axis label
plt.show()  # Showing the plot
```

4.4 Normal lines

Alongside the tangent line, there is another line that we're interested in: the normal line, a line that's perpendicular to the tangent. It's actually pretty simple to solve. By analytical geometry, if two lines are perpendicular to each other, the product of their slope must be -1.

If the slope of tangent is m_T and the normal is m_N,

$$m_T \times m_N = -1 \tag{4.62}$$

$$m_N = -\frac{1}{m_T}. \tag{4.63}$$

Since the slope of the tangent at a point $(x_0, f(x_0))$ is $\mathrm{D}_x f(x_0)$,

$$m_N = -\frac{1}{\mathrm{D}_x f(x_0)}. \tag{4.64}$$

The normal line equation can then be formed by translating the line with slope m_N to the point $(x_0, f(x_0))$:

$$y - f(x_0) = -\frac{x - x_0}{\mathrm{D}_x f(x_0)}. \tag{4.65}$$

4.5 Arc length of functions

Before we wrap off this chapter, here's a fun geometrical problem to think about. Given a function $f(x)$, what's the length of the line drawn by the function from a to b? Illustrated in fig. 4.7, the total arc length s can be subdivided into smaller arc lengths $\mathrm{d}s$ that we can sum up. How? Integrals.

The total arc length s of the function $f(x)$ from a to b is just

$$\int_a^b \mathrm{d}s \tag{4.66}$$

By the Pythagorean theorem $\mathrm{d}s = \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2}$. Can we plug this directly into the integral? Of course we can:

$$s = \int_{x=a}^{x=b} \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2}. \tag{4.67}$$

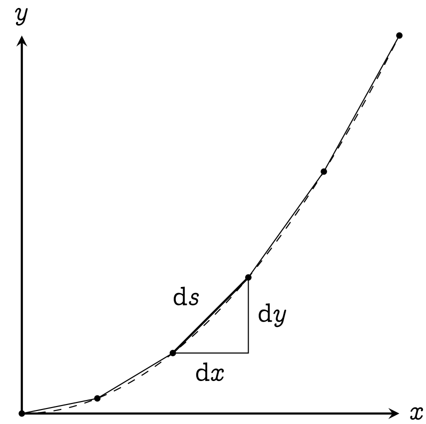


FIG. 4.7 | ARC LENGTH CALCULATION

Factorizing $\mathrm{d}x$ out of the square root would yield an integral

$$\int_{x=a}^{x=b} \sqrt{\mathrm{d}x^2 + \mathrm{d}y^2} = \int_{x=a}^{x=b} \sqrt{1 + \left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)^2} \, \mathrm{d}x \tag{4.68}$$

$$= \int_{x=a}^{x=b} \sqrt{1 + \left(\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right)^2} \, \mathrm{d}x. \tag{4.69}$$

Simply plug in the derivative of $f(x)$ and evaluate the integral, and you would get the arclength.

CHAPTER 5

Calculus, trigonometry, and oscillations

Prerequisites: *basic trigonometry.*

5.1 Invitation: oscillation and trigonometry

Trigonometry is already pretty useful on its own, as it connects the angles of a triangle to its side. But as you study further, trigonometric functions appears in many unexpected places, especially in oscillatory systems. Whether it's a motion of a spring, the synchronized blinks of fireflies, or even dynamics of love; they're somehow all modeled using trigonometric functions. Sure, the usual sine and cosine is already oscillating up and down with a definite pattern. But there are so many oscillatory functions in this world! Why these two? Is it baked into the nature of oscillation? You'll find out in this chapter.

Before moving on, I'd like to point out the whole goal of the rest of the chapters in this volume. After this chapter, the foundation of calculus are in a sense, complete. From here on out, we shall be focusing on various differential equations that describe nature; starting with this chapter

But honestly, I have no idea how I could introduce you to oscillations without studying the mathematical foundations of how trigonometric functions interact with calculus first. So, in the first few sections, it's just going to be mathematics. I'd try to keep it short.

5.1.1 Newton's fluxion notation

Here's the notation of derivative that will be used throughout the chapter, called the **Newton's fluxion notation** or, **dot notation**. *This notation is used only when the derivative is took w.r.t. time*. It places a dot over the variables, e.g., the first derivative of position r w.r.t. time is $\dot{r}$.

Higher derivatives notation is written with more dots, e.g., the second derivative of position r w.r.t. is $\ddot{r}$. The third derivative is $\dddot{r}$, forth derivative, $\ddddot{r}$ and so on.

5.2 Derivative of trigonometric functions

5.2.1 Two fundamental functions: sine and cosine

The derivative of sine oscillates along with the function itself. When $x = 0$, there's a 45° rising slope. At $x = \frac{\pi}{2}$, the slope flattens down to zero. Later at $x = \pi$, the slope is going 45° down, . If we plot the slope of sine at each point (shown as dots in fig. 5.1), it resembles a cosine curve. It might aswell be, but we'd have to mathematically prove it somehow.

Let's start with the definition of the derivative given in eq. (1.8),

$$\frac{\mathrm{d}}{\mathrm{d}x}\sin(x) = \lim_{h\to 0}\frac{\sin(x+h)-\sin(x)}{h}. \tag{5.1}$$

Using the sum of sines ($\sin(x+h) = \sin(x)\cos(h) + \cos(x)\sin(h)$),

$$\frac{\mathrm{d}}{\mathrm{d}x}\sin(x) = \lim_{h\to 0}\frac{\sin(x)\cos(h)+\cos(x)\sin(h)-\sin(x)}{h} \tag{5.2}$$

When $h \to 0$, $\cos(h) \to 1$

$$= \lim_{h\to 0}\frac{\sin(x)+\cos(x)\sin(h)-\sin(x)}{h} \tag{5.3}$$

$$= \cos(x)\lim_{h\to 0}\frac{\sin(h)}{h}. \tag{5.4}$$

For small angles, $\sin(h) \approx h$, therefore

$$= \cos(x)\lim_{h\to 0}\frac{h}{h} = \cos(x). \tag{5.5}$$

In conclusion, the derivative of sine is just the cosine just as we predicted with the graph.

Similarly, the derivative of cosine can also be found using the method of increments and using the sum of cosines formula $\cos(x+h) = \cos(x)\cos(h) - \sin(x)\sin(h)$

$$\frac{\mathrm{d}}{\mathrm{d}x}(\cos(x)) = \lim_{h\to 0}\frac{\cos(x+h)-\cos(x)}{h} \tag{5.6}$$

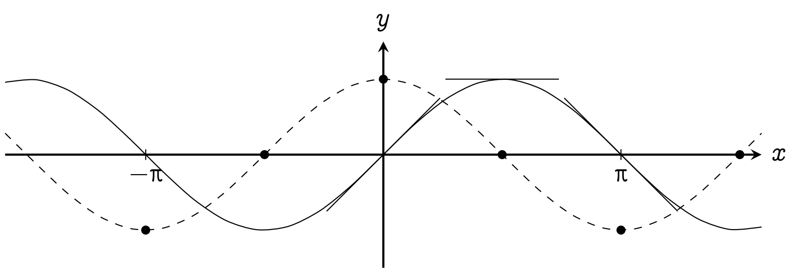


FIG. 5.1 | THE RELATION BETWEEN SINE (IN BLACK) AND COSINE (IN DASHED)

$$= \lim_{h \to 0} \frac{\cos(x)\cos(h) - \sin(x)\sin(h) - \cos(x)}{h} \tag{5.7}$$

When $h \to 0$, $\cos(h) = 1$

$$= \lim_{h \to 0} \frac{\cos(x) - \cos(x) - \sin(x)\sin(h)}{h} \tag{5.8}$$

$$= \lim_{h \to 0} \frac{-\sin(x)\sin(h)}{h} \tag{5.9}$$

For small angles of h, $\sin(h) \approx h$

$$= \lim_{h \to 0} (-\sin(x)) = -\sin(x) \tag{5.10}$$

I.e., the derivative of cosine is the negative of sine. From those two alone, we can form a set of formulas for differentiating sines and cosines:

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x}(\sin(x)) &= \cos(x) \\ \frac{\mathrm{d}^2}{\mathrm{d}x^2}(\sin(x)) &= \frac{\mathrm{d}}{\mathrm{d}x}(\cos(x)) = -\sin(x) \\ \frac{\mathrm{d}^3}{\mathrm{d}x^3}(\sin(x)) &= \frac{\mathrm{d}}{\mathrm{d}x}(-\sin(x)) = -\cos(x) \\ \frac{\mathrm{d}^4}{\mathrm{d}x^4}(\sin(x)) &= \frac{\mathrm{d}}{\mathrm{d}x}(-\cos(x)) = \sin(x) \end{aligned} \tag{5.11}$$

We can see that the derivative of these two functions cycles in four: $\sin(x), \cos(x), -\sin(x), -\cos(x)$. From these alone, we can use the rules developed in chapter 3, e.g., the chain rule, power rule, etc., in order to find the derivative of other trigonometric functions.

5.2.2 Derivative of other trigonometric functions

Note that you can continue reading the next section straightaway. This section serves as a reference. But for those who are still interested, I'm glad to have you here. This is one of the places that uses the chain rule, product rule, and power rule a lot. It's a good exercise. Without further ado, let's start with the cosecant ($\csc(x) = \frac{1}{\sin(x)}$).

$$\frac{\mathrm{d}}{\mathrm{d}x}\csc(x) = \frac{\mathrm{d}}{\mathrm{d}x}\frac{1}{\sin(x)} \tag{5.12}$$

$$= \frac{\mathrm{d}}{\mathrm{d}x}\frac{1}{u} \cdot \frac{\mathrm{d}u}{\mathrm{d}u} \qquad \text{Chain rule: } \sin(x) = u \tag{5.13}$$

$$= \frac{\mathrm{d}}{\mathrm{d}u}u^{-1} \cdot \frac{\mathrm{d}u}{\mathrm{d}x} \tag{5.14}$$

$$= -1u^{-2} \cdot \frac{\mathrm{d}}{\mathrm{d}x}\sin(x) \qquad \text{Power rule, } u = \sin(x) \tag{5.15}$$

$$= -\frac{1}{\sin^2(x)}\cos(x) \qquad \frac{\mathrm{d}}{\mathrm{d}x}\sin(x) = \cos(x) \tag{5.16}$$

$$= -\csc(x)\cot(x) \qquad \cot(x) = \frac{\cos(x)}{\sin(x)}. \tag{5.17}$$

And the secant function ($\sec(x) = \frac{1}{\cos(x)}$),

$$\frac{\mathrm{d}}{\mathrm{d}x}\sec(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{\cos(x)}\right) \tag{5.18}$$

$$= \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{u}\right) \cdot \frac{\mathrm{d}u}{\mathrm{d}u} \qquad \text{Chain rule: } \cos(x) = u \tag{5.19}$$

$$= \frac{\mathrm{d}}{\mathrm{d}u}u^{-1} \cdot \frac{\mathrm{d}u}{\mathrm{d}x} \tag{5.20}$$

$$= -1u^{-2} \cdot \frac{\mathrm{d}}{\mathrm{d}x}\cos(x) \qquad \text{Power rule, } u = \cos(x) \tag{5.21}$$

$$= -\frac{1}{\cos^2(x)}(-\sin(x)) \qquad \frac{\mathrm{d}}{\mathrm{d}x}\cos(x) = -\sin(x) \tag{5.22}$$

$$= \sec(x)\tan(x). \tag{5.23}$$

For the tangent function ($\tan(x) = \frac{\sin(x)}{\cos(x)}$), we use the product rule and the results from earlier

$$\frac{\mathrm{d}}{\mathrm{d}x}\tan(x) = \frac{\mathrm{d}}{\mathrm{d}x}\frac{\sin(x)}{\cos(x)} \tag{5.24}$$

$$= \frac{\mathrm{d}}{\mathrm{d}x}\sin(x)\sec(x) \tag{5.25}$$

$$= \sin(x)\frac{\mathrm{d}}{\mathrm{d}x}\sec(x) + \sec(x)\frac{\mathrm{d}}{\mathrm{d}x}\sin(x) \tag{5.26}$$

$$= \sin(x)\sec(x)\tan(x) + \sec(x)\cos(x) \tag{5.27}$$

$$= \sin(x)\frac{1}{\cos(x)}\frac{\sin(x)}{\cos(x)} + \frac{1}{\cos(x)}\cos(x) \tag{5.28}$$

$$= \left(\frac{\sin(x)}{\cos(x)}\right)^2 + 1 = \tan^2(x) + 1 = \sec^2(x) \tag{5.29}$$

The Pythagorean identity is used in the last line to convert $1 + \tan^2(x)$ to $\sec^2(x)$. Of course, we cannot miss its reciprocal function: $\cot(x) = \frac{1}{\tan(x)}$.

$$\frac{\mathrm{d}}{\mathrm{d}x}\cot(x) = \frac{\mathrm{d}}{\mathrm{d}x}\frac{1}{\tan(x)} \tag{5.30}$$

$$= \frac{\mathrm{d}}{\mathrm{d}x}\frac{\cos(x)}{\sin(x)} \tag{5.31}$$

$$= \frac{\mathrm{d}}{\mathrm{d}x}\cos(x)\csc(x) \tag{5.32}$$

$$= \cos(x)\frac{\mathrm{d}}{\mathrm{d}x}\csc(x) + \csc(x)\frac{\mathrm{d}}{\mathrm{d}x}\cos(x) \tag{5.33}$$

$$= \cos(x)\big(-\csc(x)\cot(x)\big) + \csc(x)\big(-\sin(x)\big) \tag{5.34}$$

$$= -\cos(x)\frac{1}{\sin(x)}\frac{\cos(x)}{\sin(x)} - \frac{1}{\sin(x)}\sin(x) \tag{5.35}$$

$$= -\big(\cot^2(x) + 1\big) = -\csc^2(x), \tag{5.36}$$

where we've again used the Pythagorean identity $\cot^2(x) + 1 = -\csc^2(x)$.

Altogether, I've summarized all the derivatives of trigonometric functions into table 5.1.

$f(x)$	$\frac{\mathrm{d}}{\mathrm{d}x}f(x)$
$\sin(x)$	$\cos(x)$
$\cos(x)$	$-\sin(x)$
$\tan(x)$	$\sec^2(x)$
$\csc(x)$	$-\csc(x)\cot(x)$
$\sec(x)$	$\sec(x)\tan(x)$
$\cot(x)$	$-\csc^2(x)$

TABLE 5.1 | THE DERIVATIVES OF TRIGONOMETRIC FUNCTIONS

5.3 Power series of trigonometric functions

The approximation method that I'd be discussing are commonly taught to high schoolers with the name **small angle approximation**, i.e., for $\theta \to 0$,

$$\sin(\theta) \approx \tan(\theta) \approx \theta \quad \text{and,} \quad \cos(\theta) \approx 1 - \frac{\theta^2}{2} \tag{5.37}$$

These approximations are very useful for calculating limits, e.g.,[1]

$$\lim_{\theta \to 0} \frac{\sin(n\theta)}{\theta} = \lim_{\theta \to 0} \frac{n\theta}{\theta} = n. \tag{5.38}$$

Here, I want to focus on where these approximations really comes from: the power series development of trigonometric functions.

5.3.1 Power series of sine

Assume that the sine function can be written as a power series

$$\sin(x) = s_0 + s_1 x^1 + s_2 x^2 + s_3 x^3 + \ldots \tag{5.39}$$

Because $\sin(0) = 0$,

$$\sin(0) = s_0 + s_1(0)^1 + s_2(0)^2 + s_3(0)^3 + \ldots \tag{5.40}$$

$$s_0 = 0; \tag{5.41}$$

Because sine is an odd function, $\sin(-x) = -\sin(x)$; thus,

$$\begin{aligned} &s_1(-x)^1 + s_2(-x)^2 \\ &\quad + s_3(-x)^3 + s_4(-x)^4 + \ldots = -s_1 x^1 - s_2 x^2 \\ &\qquad\qquad - s_3 x^3 - s_4 x^4 + \ldots \end{aligned} \tag{5.42}$$

$$\begin{aligned} &-s_1 x^1 + s_2 x^2 \\ &\quad - s_3 x^3 + s_4 x^4 - \ldots = -s_1 x^1 - s_2 x^2 - \\ &\qquad\qquad s_3 x^3 - s_4 x^4 - \ldots \end{aligned} \tag{5.43}$$

[1] Normally these limits are evaluated using the squeeze theorem. However, I don't want the mathematical intricacies to disrupt the flow of our journey right now. The full theorem will be discussed later at the end of the book.

In the L.H.S., the sign alternates between negative and positive. In order for the L.H.S. to match the R.H.S., all the positive terms must vanish, leaving only the matching negative terms; therefore,

$$\sin(x) = s_1 x^1 + s_3 x^3 + s_5 x^5 + \ldots . \tag{5.44}$$

This is what we mean by "sine is an odd function": there are only odd powers in its power series.

To find s_1, take the derivative of sine

$$\frac{\mathrm{d}}{\mathrm{d}x}\sin(x) = \cos(x) = s_1 + 3s_3x^2 + 5s_5x^4 + \ldots , \tag{5.45}$$

then substitute $x = 0$

$$\cos(0) = 1 = s_1 + 3s_3(0)^2 + 5s_5(0)^4 + \ldots \tag{5.46}$$

$$1 = s_1. \tag{5.47}$$

We can then take the derivative of sine again to get a recurrence relation.

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}\sin(x) = 3\cdot 2s_3x + 5\cdot 4s_5x^3 + 7\cdot 6s_7x^5 + \ldots \tag{5.48}$$

$$-\sin(x) = \frac{3!}{1!}s_3x + \frac{5!}{3!}s_5x^3 + \frac{7!}{5!}s_7x^5 + \ldots \tag{5.49}$$

$$-s_1x - s_3x^3 - s_5x_5 - \ldots = \frac{3!}{1!}s_3x + \frac{5!}{3!}s_5x^3 + \frac{7!}{5!}s_7x^5 + \ldots \tag{5.50}$$

By matching terms with the same order,

$$\begin{aligned} -s_1 &= \frac{3!}{1!}s_3x, \\ -s_3 &= \frac{5!}{3!}s_5x, \\ -s_5 &= \frac{7!}{5!}s_7x. \\ &\vdots \end{aligned} \tag{5.51}$$

From $s_1 = 1$, we get that $s_3 = -\frac{1}{3!}$, $s_5 = -\frac{1}{5!}$, $s_7 = -\frac{1}{7!}$ etc. The negative sign alternates every term. We can then summarize the whole thing as

$$\sin(x) = \sum_{n=0}^{\infty} \frac{(-1)^{2n+1}}{(2n+1)!}x^{2n+1} \tag{5.52}$$

$$= x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \frac{1}{7!}x^7 + \ldots, \tag{5.53}$$

which is the polynomial expansion of the sine function.

5.3.2 Power series of cosine

The power series of cosine can be developed using the same method as discussed earlier. But notice that the derivative of sine is cosine. Hence, we can just take the derivative of the sine's power series once to get the cosine's power series.

$$\frac{\mathrm{d}}{\mathrm{d}x}\sin(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left(x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \frac{1}{7!}x^7\right) \tag{5.54}$$

$$\cos(x) = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \frac{1}{6!}x^6 + \cdots = \sum_{n=0}^{\infty}\frac{(-1)^{2n}}{(2n)!}x^{2n} \tag{5.55}$$

5.3.3 Power series of other trigonometric functions

The same method could be used to find the power series expansion of secant, cosecant, tangent, etc. However, their power series doesn't actually represent the function because the other functions are all ratios of sine and cosines. To see what I mean, consider the behavior of the tangent function.

The tangent function is defined as $\tan\theta = \frac{\sin\theta}{\cos\theta}$. From this definition, $\tan\theta$ approaches infinity whenever $\cos\theta$ approaches zero, which is at $n\pi$ where $n \in \mathbb{Z}$[2]. As seen, there are infinitely many places in which the function approaches infinity. However, polynomials only have two infinities: at $x \to \infty$, and $x \to -\infty$; therefore, any polynomial expansion can't possibly represent the tangent function. The same logic also works for other trigonometric functions.

[2] $\mathbb{Z}$ is the set of integers

5.3.4 Approximation of trigonometric functions

One way a function can be approximated is by truncating its power series. To see why this is viable, consider the limit

$$\lim_{x\to 0}\sin(x) = \lim_{x\to 0}\left(x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \frac{1}{7!}x^7 + ...\right) \tag{5.56}$$

When $x \to 0$, all the higher order degrees term vanishes; thus,

$$\lim_{x\to 0}\sin(x) = x. \tag{5.57}$$

It tells us that when $x \approx 0$, the sine function behaves like a linear function. Illustrated in fig. 5.2,

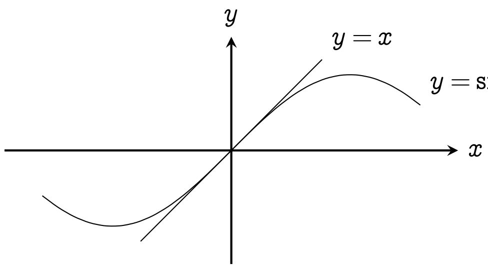


FIG. 5.2 | COMPARISON OF $y = \sin(x)$ AND $y = x$ NEAR $x = 0$

When x gets larger, the higher order terms in the power series contributes more and more, so we need to include them. An excellent approximation for the sine function when $x \in [-\pi, \pi]$ is a truncation at the fifth term:

$$\sin(x) \approx x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \frac{1}{7!}x^7 + \frac{1}{9!}x^9, \tag{5.58}$$

plotted in fig. 5.3. It is extremely accurate in that range. For cosine,

$$\lim_{x\to 0}\cos(x) = \lim_{x\to 0}\left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - ...\right) \approx 1 - \frac{x^2}{2!} \approx 1. \tag{5.59}$$

An acceptable approximation around zero is 1 or $1 - \frac{x^2}{2}$. Similarly, an excellent approximation when $x \in [-\pi, \pi]$ is also a truncation at the fourth

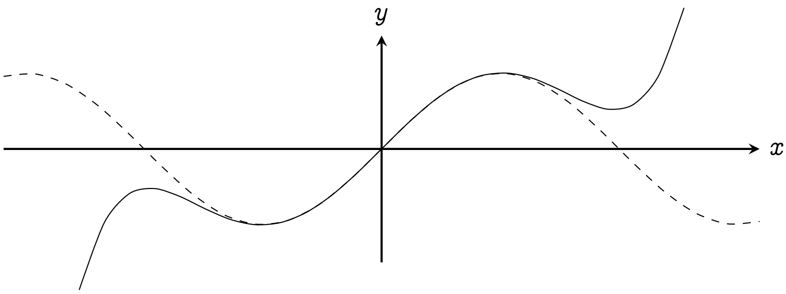


FIG. 5.3 | AN APPROXIMATION OF SINE BY TRUNCATING ITS POWER SERIES AT THE FIFTH TERM.

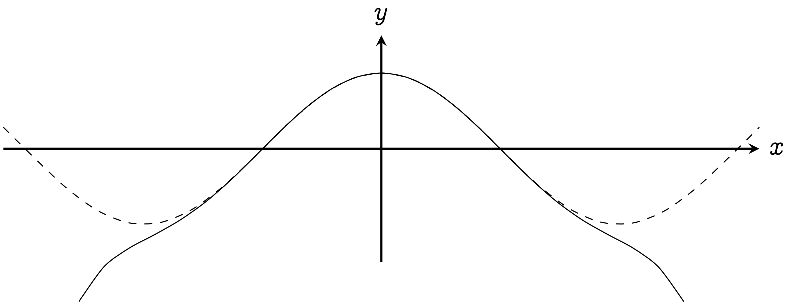


FIG. 5.4 | AN APPROXIMATION OF COSINE BY TRUNCATING ITS POWER SERIES AT THE FOURTH TERM

term:

$$\cos(x) \approx 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \frac{1}{6!}x^6 \tag{5.60}$$

plotted in fig. 5.4

5.4 The harmonic oscillator

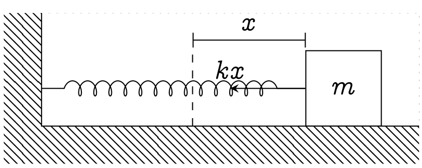


FIG. 5.5 | SETUP OF THE HARMONIC OSCILLATOR

Finally, we're ready to tackle our first oscillatory system: the harmonic oscillator. Illustrated in fig. 5.5, consider a mass m connected to a spring. The point that the spring is neither stretched or compressed is called the **equilibrium position**, drawn as the dashed line. The force that the spring acts on the mass depends on the displacement from the equilibrium position x:

$$F = -kx, \tag{5.61}$$

where k is the stiffness of the spring. Higher k represents a stiffer spring, and lower k represents a looser spring. Newton's second law reads

$$-kx = m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} \tag{5.62}$$

$$-\frac{k}{m}x = \frac{\mathrm{d}^2x}{\mathrm{d}t^2}. \tag{5.63}$$

We're interested in the solution of eq. (5.63) given an initial condition that at time $t = 0$, the mass is distance A away from the equilibrium position, and it's released with zero initial velocity, i.e.,

$$x(t=0) = A \quad \text{and,} \quad \dot{x}(t=0) = 0. \tag{5.64}$$

5.4.1 Your first trigonometric substitution

The equation that we got, eq. (5.63), is an inseparable second order differential equation. It is extremely hard to work with. But by exploiting the chain rule, we can reduce it into a first order differential equation.

$$\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right) = \frac{\mathrm{d}\dot{x}}{\mathrm{d}t} = \frac{\mathrm{d}\dot{x}}{\mathrm{d}x}\frac{\mathrm{d}x}{\mathrm{d}t} = \dot{x}\frac{\mathrm{d}\dot{x}}{\mathrm{d}x}. \tag{5.65}$$

Equation (5.63) then becomes

$$-\frac{k}{m}x = \dot{x}\frac{\mathrm{d}\dot{x}}{\mathrm{d}x} \tag{5.66}$$

$$-\frac{k}{m}x\,\mathrm{d}x = \dot{x}\,\mathrm{d}\dot{x} \tag{5.67}$$

$$-\frac{k}{m}\int x\,\mathrm{d}x = \int \dot{x}\,\mathrm{d}\dot{x} \tag{5.68}$$

$$-\frac{k}{m}\frac{x^2}{2} = \frac{\dot{x}^2}{2} + C \tag{5.69}$$

$$-\frac{k}{m}x^2 = \dot{x}^2 + C. \tag{5.70}$$

To find C, substitute in the initial condition eq. (5.64).

$$-\frac{k}{m}(A)^2 = (0)^2 + C \tag{5.71}$$

$$C = -\frac{k}{m}A^2. \tag{5.72}$$

Thus,

$$-\frac{k}{m}x^2 = \dot{x}^2 - \frac{k}{m}A^2 \tag{5.73}$$

$$\dot{x}^2 = \frac{k}{m}\left(A^2 - x^2\right) \tag{5.74}$$

$$\dot{x} = \sqrt{\frac{k}{m}}\sqrt{A^2 - x^2}. \tag{5.75}$$

Because I'm too lazy to write square roots, let's use $\omega = \sqrt{k\,/\,m}$; thus,

$$\dot{x} = \omega\sqrt{A^2 - x^2} \tag{5.76}$$

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \omega\sqrt{A^2 - x^2} \tag{5.77}$$

$$\omega\,\mathrm{d}t = \frac{1}{\sqrt{A^2 - x^2}}\,\mathrm{d}x \tag{5.78}$$

$$\omega\int \mathrm{d}t = \int \frac{1}{\sqrt{A^2 - x^2}}\,\mathrm{d}x. \tag{5.79}$$

The L.H.S. directly evaluates to ωt. But on the R.H.S., power rule substitution wouldn't work because there's a polynomial with degree two the root, but no terms with lower degree outside. If $u = A^2 - x^2$, we'd be going in loops. So how do we evaluate this?

The answer is, you can't do anything about it if we're going to use the methods so far. It's time for a new tool: **trigonometric substitution**.

The idea is very simple. We want to take advantage of the Pythagorean identity $1 - \sin^2\theta = \cos^2\theta$ to bring $A^2 - x^2$ out of the root. Substitute in $x = A\sin\theta$ and see what happens.

$$\int \frac{1}{\sqrt{A^2 - x^2}}\,\mathrm{d}x = \int \frac{1}{\sqrt{A^2 - A^2\sin^2\theta}}\,\mathrm{d}x \tag{5.80}$$

$$= \frac{1}{A}\int \frac{1}{\sqrt{1 - \sin^2\theta}}\,\mathrm{d}x \tag{5.81}$$

$$= \frac{1}{A}\int \frac{1}{\sqrt{\cos^2\theta}}\,\mathrm{d}x = \frac{1}{A}\int \frac{1}{\cos\theta}\,\mathrm{d}x \tag{5.82}$$

Taking care of the $\mathrm{d}x$; if $x = A\sin\theta$, then

$$\frac{\mathrm{d}x}{\mathrm{d}\theta} = A\frac{\mathrm{d}\sin\theta}{\mathrm{d}\theta} \tag{5.83}$$

$$\frac{\mathrm{d}x}{\mathrm{d}\theta} = A\cos\theta \tag{5.84}$$

$$\mathrm{d}x = A\cos\theta\,\mathrm{d}\theta. \tag{5.85}$$

Substituting back into the integral gives

$$\frac{1}{A}\int \frac{1}{\cos\theta}\big(A\cos\theta\,\mathrm{d}\theta\big) = \int \mathrm{d}\theta = \theta + C. \tag{5.86}$$

Because $x = A\sin\theta$, $\theta = \arcsin\big(x\,/\,A\big)$

$$\int \frac{1}{\sqrt{A^2 - x^2}}\,\mathrm{d}x = \arcsin\left(\frac{x}{A}\right) + C. \tag{5.87}$$

Thus, eq. (5.79) becomes

$$\omega t = \arcsin\left(\frac{x}{A}\right) + C. \tag{5.88}$$

Then, substitute in the initial condition (eq. (5.64)) to find C:

$$\omega(0) = \arcsin\left(\frac{A}{A}\right) + C \tag{5.89}$$

$$C = \frac{\pi}{2}. \tag{5.90}$$

Thus,

$$\omega t = \arcsin\left(\frac{x}{A}\right) + \frac{\pi}{2} \tag{5.91}$$

$$\arcsin\left(\frac{x}{A}\right) = \omega t + \frac{\pi}{2} \tag{5.92}$$

$$x = A\sin\left(\omega t + \frac{\pi}{2}\right). \tag{5.93}$$

Since $\sin(\theta + \frac{\pi}{2}) = \cos(\theta)$,

$$x(t) = A\cos(\omega t) \tag{5.94}$$

which is the solution of the differential equation. It represents where the mass is relative to the equilibrium point over time. Taking the derivative of this once,

$$\dot{x}(t) = A\frac{\mathrm{d}\cos(\omega t)}{\mathrm{d}t} \tag{5.95}$$

Let $u = \omega t$. By the chain rule,

$$= A\frac{\mathrm{d}\cos(u)}{\mathrm{d}u} \times \frac{\mathrm{d}u}{\mathrm{d}t} \tag{5.96}$$

$$= -A\sin(u)\frac{\mathrm{d}}{\mathrm{d}t}(\omega t) \tag{5.97}$$

$$= -\omega A\sin(\omega t), \tag{5.98}$$

we get the velocity as a function w.r.t. time. In a similar manner, taking the derivative of velocity again gives the acceleration

$$\ddot{x} = \frac{\mathrm{d}\dot{x}}{\mathrm{d}t} = \frac{\mathrm{d} - A\omega\sin(\omega t)}{\mathrm{d}t} = -\omega^2 A\sin(\omega t). \tag{5.99}$$

5.4.2 The power series method

Perhaps a faster method to solve eq. (5.63) is to just assume that the solution can be expanded in terms of power series

$$x(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3 + \dots . \tag{5.100}$$

Consequently,

$$\dot{x}(t) = a_1 + 2a_2 t + 3a_3 t^2 + 4a^4 t_3 + \dots \tag{5.101}$$

$$\ddot{x}(t) = 2a_2 + 3 \cdot 2a_3 t + 4 \cdot 3a_4 t^2 + 5 \cdot 4a_5 t_3 + ... \tag{5.102}$$

The initial condition given in eq. (5.64) ($x(t = 0) = A$, $\dot{x}(t = 0) = 0$) can then be used to find a_0 and a_1.

$$x(t = 0) = a_0 + a_1(0) + a_2(0)^2 + a_3(0)^3 + ... \tag{5.103}$$

$$a_0 = A \tag{5.104}$$

and from eq. (5.101),

$$\dot{x}(t = 0) = a_1 + 2a_2(0) + 3a_3(0)^2 + 4a_4(0)^3 + ... \tag{5.105}$$

$$a_1 = 0 \tag{5.106}$$

The rest of the terms can be found by using the differential equation itself. First, rearrange the equation into

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} + \omega^2 x = 0. \tag{5.107}$$

Then, substitute in eqs. (5.100) and (5.102)

$$0 = \frac{\mathrm{d}^2 x}{\mathrm{d}t^2} + \omega^2 x \tag{5.108}$$

$$\begin{aligned} 0 = &\left(2a_2 + 3 \cdot 2a_3 t + 4 \cdot 3a_4 t^2 + 5 \cdot 4a_5 t^3 + 6 \cdot 5a_6 t^4 + 6 \cdot 7a_7 t^5 + ...\right) \\ &+ \left(\omega^2 A + a_2\omega^2 t^2 + a_3\omega^2 t^3 + a_4\omega^2 t^4 + a_5\omega^2 t^5 ...\right) \end{aligned}$$

$$\begin{aligned} 0 = &\left(2a_2 + \omega^2 A\right) \\ &+ (3 \cdot 2a_3)t \\ &+ (4 \cdot 3a_4 + \omega^2 a_2)t^2 \\ &+ (5 \cdot 4a_5 + \omega^2 a_3)t^3 \\ &+ (6 \cdot 5a_6 + \omega^2 a_4)t^4 \\ &+ (7 \cdot 6a_7 + \omega^2 a_5)t^5 + ... \end{aligned} \tag{5.109}$$

For the R.H.S. to be 0, all terms must vanish; thus, we get an infinite series of equations to be evaluated

$$\begin{aligned}
0 &= \left(2a_2 + \omega^2 A\right) \\
0 &= (3 \cdot 2a_3) \\
0 &= (4 \cdot 3a_4 + \omega^2 a_2) \\
0 &= (5 \cdot 4a_5 + \omega^2 a_3) \\
0 &= (6 \cdot 5a_6 + \omega^2 a_4) \\
0 &= (7 \cdot 6a_7 + \omega^2 a_5) \\
&\vdots
\end{aligned} \tag{5.110}$$

Starting from the second one: $0 = 3 \cdot 2a_3$ means that $a_3 = 0$. This eliminates every term that includes a_3. The fourth equation says $5 \cdot 4a_5 + \omega^2 a_3 = 0$, thus $a_5 = 0$. This pattern continues, eliminating every odd terms in the series. We're then left with

$$\begin{aligned}
0 &= 2a_2 + \omega^2 A \\
0 &= 4 \cdot 3a_4 + \omega^2 a_2 \\
0 &= 6 \cdot 5a_6 + \omega^2 a_4 \\
0 &= 8 \cdot 7a_8 + \omega^2 a_6 \\
&\vdots
\end{aligned} \tag{5.111}$$

Starting from the first one, we get $a_2 = -\frac{\omega^2}{2} A$. The second equation gives $a_4 = \frac{\omega^4}{4!} A$. The third gives $a_6 = \frac{\omega^6}{6!} A$. This pattern of $a_{2n} = \frac{\omega^{2n}}{2n} A$ continues on forever. The power series expansion for $x(t)$ gives

$$x(t) = a_0 + \cancelto{0}{a_1} + a_2 t^2 + \cancelto{0}{a_3} t^3 + a_4 t^4 + \cancelto{0}{a_5} t^5 + a_6 t^6 + \ldots \tag{5.112}$$

$$= A + A\frac{\omega^2}{2!} t^2 + A\frac{\omega^4}{4!} t^4 + A\frac{\omega^6}{6!} t^6 + \ldots \tag{5.113}$$

$$= A\left(\frac{(\omega t)^2}{2!} + \frac{(\omega t)^4}{4!} + \frac{(\omega t)^6}{6!} + \ldots\right) \tag{5.114}$$

Looks familiar yet? *This*, the result that we got is exactly the power series expansion for cosine (eq. (5.55)); thus,

$$x(t) = A\cos(\omega t). \tag{5.115}$$

5.4.3 Physical aspects of the harmonic oscillator

Now that we have solved the harmonic oscillator, let's bring our attention to the qualitative description of the system.

The mass is released with zero velocity at $x = A$. As the spring tries to restore its equilibrium, it accelerates the mass towards the equilibrium position. The further the mass is from the equilibrium, the stronger the spring pulls. Initially, the spring pulls very hard on the mass, causing it to accumulate much velocity. But as it approaches the equilibrium position, the spring's force diminishes and there is barely any force left to stop the mass, so it overshoots the other side, compressing the spring. This oscillatory cycle continues on forever. Intuitively, we can predict that the mass will have the highest velocity and lowest acceleration as it directly passes through the equilibrium; and, lowest velocity and highest acceleration when the spring is fully stretched or compressed. But does the mathematics agree with this?

In fig. 5.6, the position function (eq. (5.94)) is plotted in thick lines, and the velocity function (eq. (5.98)) are plotted in dashed lines.

The furthest point that a mass could be from a spring is at $x = \pm A$. Mathematically, this represents the local maxima and minima of the function, which can be obtained by taking the first derivative and setting it to zero (section 4.2). The derivative of the position w.r.t. time is just the velocity (eq. (5.98)), and it reaches zero whenever $\sin(\omega t)$ is zero

$$\sin(\omega t) = 0 \tag{5.116}$$

$$\omega t = \arcsin(0) + 2n\pi \qquad n \in \mathbb{Z} \tag{5.117}$$

$$t = \frac{(2n+1)\pi}{\omega}. \tag{5.118}$$

Another way to derive this fact is to just notice that the range of the cosine function only extends from -1 to 1; therefore, the highest that $A\cos(\omega t)$ can go is just A. So, another equation that we can form is

$$\cos(\omega t) = 1 \tag{5.119}$$

$$\omega t = \arccos(1) + 2n\pi \qquad n \in \mathbb{Z} \tag{5.120}$$

$$t = \frac{(2n+1)\pi}{\omega}, \tag{5.121}$$

which has the same solution as $\sin(\omega t) = 0$.

The time that the mass goes through the equilibrium position can be found by just setting $A\cos(\omega t)$ to zero

$$A\cos(\omega t) = 0 \tag{5.122}$$

$$\omega t = \arccos(0) + 2n\pi \qquad n \in \mathbb{Z} \tag{5.123}$$

$$t = \frac{2n\pi}{\omega} \tag{5.124}$$

Since the velocity function (eq. (5.98)) is a sine function, it must oscillates π radians off-phase relative to the position function. Therefore,

1. Whenever the mass is furthest away from the equilibrium, the velocity is zero
2. Whenever the mass is directly passing through the equilibrium, the velocity reaches its maximum at $\pm\omega A$

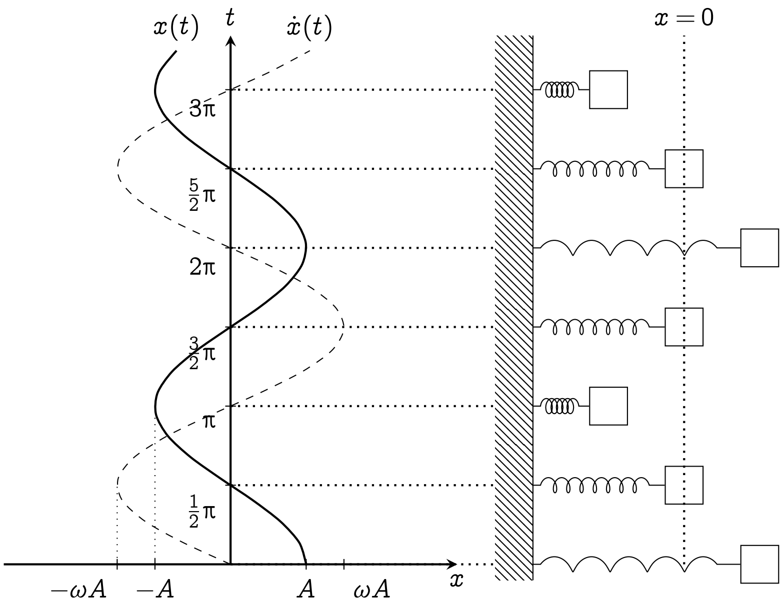


FIG. 5.6 | THE TRAJECTORY OF A MASS IN THE HARMONIC OSCILLATOR SYSTEM. (LEFT) THE POSITION IS PLOTTED IN THICK LINES, AND THE VELOCITY, IN DOTTED LINES. (RIGHT) DEPICTION OF AN OSCILLATING MASS, ACCORDING TO THE SOLUTION OF THE DIFFERENTIAL EQUATION.

The first fact can be easily verified by setting the velocity function, $-\omega A \sin(\omega t)$ equals to zero

$$-\omega A \sin(\omega t) = 0 \tag{5.125}$$

$$\omega t = \arcsin(0) + 2n\pi \tag{5.126}$$

$$t = \frac{(2n+1)}{\omega}, \tag{5.127}$$

which directly matches with the point of furthest position (eq. (5.118)). The second one can be verified by taking the derivative of velocity, which is just the acceleration, found in eq. (5.99). Setting that equals to zero yields

$$-\omega^2 A \cos(\omega t) = 0 \tag{5.128}$$

$$\omega t = \arccos(0) + 2n\pi \tag{5.129}$$

$$t = \frac{2n\pi}{\omega}, \tag{5.130}$$

which matches our result from eq. (5.124).

The acceleration function is again, a cosine function that's in phase with the position. Therefore, the local minima and maxima from the position function must match. We can quickly conclude that the acceleration is at its highest at $t = \frac{2n\pi}{\omega}$, and is zero at $\frac{(2n+1)\pi}{\omega}$.

You would think that we have completed the physical aspects of the harmonic oscillator. But there is something I left over: the ω term. Since the beginning, I have assigned $\omega = \sqrt{\frac{k}{m}}$ as a notational trick for convenience. Is there any real physical meaning behind ω?

If you don't know much about units, it's fine if you skip these paragraphs straightaway. But one clue that suggests the physical meaning of ω is its unit. Since $\omega = \sqrt{\frac{k}{m}}$, we have to analyse the units of its components separately. k is a spring stiffness constant which is got from

$$F = -kx \quad \text{or,} \quad k = -\frac{F}{x}. \tag{5.131}$$

F is the unit of force (Newton N), which is just $\mathrm{kg}\frac{\mathrm{m}}{\mathrm{s}^2}$), and x, the units of distance (Meters m); therefore, k is the unit of

$$\frac{\mathrm{kg}\frac{\mathrm{m}}{\mathrm{s}^2}}{\mathrm{m}} = \frac{\mathrm{kg}}{\mathrm{s}^2}. \tag{5.132}$$

And since m has the unit of kilogram (kg), ω which is $\sqrt{\frac{k}{m}}$ has the unit of

$$\sqrt{\frac{\frac{\mathrm{kg}}{\mathrm{s}^2}}{\mathrm{kg}}} = \sqrt{\frac{1}{\mathrm{s}^2}} = \frac{1}{\mathrm{s}} \tag{5.133}$$

Funnily enough, this is the unit of Hertz, ($\frac{1}{\mathrm{s}} = \mathrm{Hz}$) which describes a frequencies. But it cannot possibly be just a unit of frequency because the solution of the differential equation says so.

The solution of our equation is $x(t) = A\cos(\omega t)$. If ω is really a unit of Hertz, then ωt would have no units, which contradicts that the cosine function either accepts radians (rad), or degrees (deg). But these are units that are dimensionless, meaning that we can just multiply it into ω; therefore, ω must have the unit of $\mathrm{rad\,s^{-1}}$, which is the unit of angular frequencies.

The conclusion that ω is the angular frequencies can also be directly deducted from the solution of the equation. From both of eq. (5.118) and eq. (5.124), the time between each maximum and minima point is divided by ω. If ω is very high, then the time between the maximum and minimum is decreased. If it's low, then the time is increased. Therefore, ω is directly associated to the natural frequency of the system. The more ω is, the more frequent the oscillator oscillates.

Another way ω can be interpretted is by looking at the expression of ω itself: $\sqrt{\frac{k}{m}}$. The spring stiffness constant k is on the top, so it means the stiffer the spring, the higher omega is. A stiffer spring would allow the spring to pull the mass faster; thus, increasing the frequencies. The mass m is on the bottom: the larger the mass, the lower the omega. That just means a more massive block is harder for the spring to move; therefore, the mass will travel slower, decreasing the frequencies.

5.4.4 Solution for other initial conditions. A trick with simplifying inverse trigonometric functions.

Now, let's see what happens if the initial condition is changed from $x(t=0) = A$, and $v(t=0) = 0$, to

$$x(t=0) = X \quad \text{and,} \quad v(t=0) = V. \tag{5.134}$$

Since it's the same system but just under a different initial condition, the differential equation stays exactly the same

$$\frac{\mathrm{d}^2 x}{\mathrm{d}t^2} + \omega^2 x = 0, \tag{5.135}$$

where $\omega = \sqrt{k \,/\, m}$ for the same notational reason. In a similar manner, use the chain rule on the second derivative to reduce it down to a first order differential equation:

$$\dot{x}\frac{\mathrm{d}\dot{x}}{\mathrm{d}x} = -\omega^2 x \tag{5.136}$$

$$\dot{x}^2 = -\omega^2 x^2 + C. \tag{5.137}$$

To find the initial condition, set $t = 0$ and substitute $x(t = 0) = X$ and $v(t = 0) = V$.

$$V^2 = -\omega^2 (X)^2 + C \tag{5.138}$$

$$C = \omega^2 X^2 + V^2. \tag{5.139}$$

Since the form is quite complex, let's call this C_1, and substitute it in later when the differential equation is fully solved. So,

$$\dot{x}^2 = -\omega^2 x^2 + C_1 \tag{5.140}$$

$$\frac{\mathrm{d}x}{\mathrm{d}t} = \sqrt{-\omega^2 x^2 + C_1} \tag{5.141}$$

$$\int \mathrm{d}t = \int \frac{1}{\sqrt{C_1 - \omega^2 x^2}}\,\mathrm{d}x \tag{5.142}$$

$$t = \int \frac{1}{\frac{1}{\omega}\sqrt{\left(\sqrt{C_1} \,/\, \omega\right)^2 - x^2}}\,\mathrm{d}x \tag{5.143}$$

$$t = \omega \int \frac{1}{\sqrt{\left(\sqrt{C_1} \,/\, \omega\right)^2 - x^2}}\,\mathrm{d}x \tag{5.144}$$

Since C_1 is always positive due to its term all being squared. The same trigonometric substitution can be used to obtain

$$t = \omega \arcsin\left(\frac{x}{\sqrt{C_1} \,/\, \omega}\right) + C_2 \tag{5.145}$$

Then, set $t = 0$, and substitute $x(t = 0) = X$ to find C_2.

$$0 = \omega \arcsin\left(\frac{\omega X}{\sqrt{C_1}}\right) + C_2 \tag{5.146}$$

$$C_2 = -\omega \arcsin\left(\frac{\omega X}{\sqrt{\omega^2 X^2 + V^2}}\right) \tag{5.147}$$

5.4.5 The ansatz method

5.5 Damped harmonic oscillator

5.6 Substitution to eliminate trigonometric functions

5.7 Circles and ellipses

5.7.1 Perimeter

5.7.2 Areas of circle. A funny way to approximate pi.

5.8 Volume of revolutions

CHAPTER 6

Merging calculus and dynamical systems

As said earlier, with the merge of calculus and trigonometry, all the cards are on the deck. This is all the calculus knowledge there is from here to the nineteenth-century.

6.1 Further examples of calculus in physics and kinematics

6.1.1 Block sliding down a ramp with friction

Illustrated in fig. 6.1, there are two axis of motion: perpendicular and parallel to the block's expected motion. From what we know, there are the gravity mg that pulls the block straight down and the friction force f_r acting against the block's expected motion in the parallel axis. The problem constraints use along the perpendicular axis: the block can't move

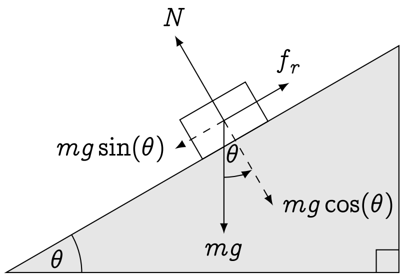


Fig. 6.1 | A block sliding down a rough ramp (i.e. with friction) angled θ relative to the ground

along that perpendicular axis *because there's a ramp in the way*. Therefore, the ramp must act a force N on the block.

We want to find the equation of motion for this block that's sliding down a ramp with friction. As far as the perpendicular axis goes, we don't have to worry about that one because nothing is moving there anyways. On the parallel axis, there's the f_r and $mg\sin(\theta)$.[1] If we set the direction of the parallel axis to pointing downwards along the block's movement, we get that the total force is

$$F_n = mg\sin(\theta) - f_r\text{[2]}.$$

By the Newton's second law,

$$\begin{aligned} m\frac{\mathrm{d}t}{\mathrm{d}(}\frac{\mathrm{d}x}{\mathrm{d}t}) &= mg\sin(\theta) - f_r \\ \frac{\mathrm{d}t}{\mathrm{d}(}\frac{\mathrm{d}x}{\mathrm{d}t}) &= g\sin(\theta) - \frac{f_r}{m}. \end{aligned}$$

$g\sin(\theta) - f_r\ /\ m$ doesn't change with time, let's name it κ. The equation then becomes

$$\frac{\mathrm{d}t}{\mathrm{d}(}\frac{\mathrm{d}x}{\mathrm{d}t}) = \kappa \tag{6.1}$$

[1] Just decompose mg into its parallel and perpendicular axis. Using basic trigonometry is enough.

[2] F_n for F_{net} or, total force.

$$\begin{aligned}
\int \mathrm{d}(\frac{\mathrm{d}x}{\mathrm{d}t}) &= \int \kappa \,\mathrm{d}t \\
\frac{\mathrm{d}x}{\mathrm{d}t} &= \kappa t \\
\int \mathrm{d}x &= \int \kappa t \,\mathrm{d}t \\
x(t) &= \left(g \sin(\theta) - \frac{f_r}{m}\right) \frac{t^2}{2}. \qquad (6.2)
\end{aligned}$$

which *is* our equation of motion. Notice, eq. (6.1) literally has the same form as **??** that we derived from "ball dropped from a building". And indeed, it should be the same because it's just a thing that's under a constant acceleration.

6.1.2 One-dimensional movement with drag forces

The free body diagram is illustrated in **??**. There's the gravity $m\mathbf{g}$ pulling the ball down, and drag force $kf_r(\mathbf{v})$. However, drag is a complex thing. There is no such thing as an "exact drag function" because drag depends on so many variables, e.g., air viscosity, air compressibility, object's shape, surface's friction, just to name a few. Therefore, the drag function $f_r(\mathbf{v})$ is a *simplified model*, not the real thing.

We shall model the drag based on two assumptions. i.) the drag force should depend on the velocity : the faster, the more drag. And, ii.) any function can be approximated using the power series expansion (also discussed in **??**):

$$f_r(\mathbf{v}) = a_0 + a_1\mathbf{v} + a_2\mathbf{v}^2 + a_3\mathbf{v}^3 + \dots .$$

Considering only the first three terms should be enough. We know that a_0 must be 0, because otherwise our object would just accelerate all the time, which is no good. Therefore, there can only be $a_1\mathbf{v}+a_2\mathbf{v}^2$. Newton's second law reads

$$m\frac{\mathrm{d}}{\mathrm{d}t}(\frac{\mathrm{d}x}{\mathrm{d}t}) = a_1\frac{\mathrm{d}x}{\mathrm{d}t} + a_2\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2.$$

Using fluxion notation,

$$\begin{aligned} m\frac{\mathrm{d}\dot{x}}{\mathrm{d}t} &= a_1\dot{x} + a_2\dot{x}^2 \\ \frac{\mathrm{d}\dot{x}}{\mathrm{d}t} &= \frac{a_1}{m}\dot{x} + \frac{a_2}{m}\dot{x}^2 \end{aligned}$$

For convenience, let $a_1 \,/\, m = p$ and $a_2 \,/\, m = q$. The equation reads

$$\frac{\mathrm{d}\dot{x}}{\mathrm{d}t} = p\dot{x} + q\dot{x}^2. \tag{6.3}$$

It is obvious that both p and q must be negative, otherwise the object would accelerate forward with the velocity. Frankly, eq. (6.3) is not possible to solve using the techniques that we have now. I'll revisit this exact differential equation later in **??**. For now, we shall deal with a simpler equation by considering two cases: only linear drag and only quadratic drag.

Motion with just linear drag

You don't really see linear drag in real life. It's mostly drag in moving liquid, e.g., a fish swimming in the water. Equation (6.3) simplifies to

$$\frac{\mathrm{d}\dot{x}}{\mathrm{d}t} = p\dot{x}.$$

I shall set v_0 as the initial velocity and x_0 as the initial position. There are two methods of solving this. First, by separating variables using analytical methods.

$$\begin{aligned} \int \frac{1}{\dot{x}}\,\mathrm{d}\dot{x} &= p\int \mathrm{d}t \\ \ln(\dot{x}) + C &= pt. \end{aligned} \tag{6.4}$$

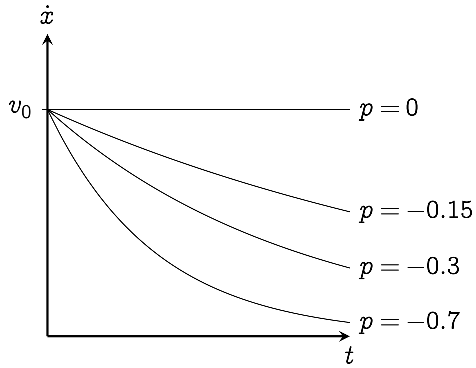


FIG. 6.2 | PLOT OF EQ. (6.5) WITH VARIOUS p.

To find out what this integration constant should be, we have to use the initial condition $t = 0 \implies \dot{x} = v_0$.

$$\begin{aligned} \ln(v_0) + C &= 0 \\ C &= -\ln(v_0). \end{aligned}$$

Equation (6.4) then becomes

$$\begin{aligned} \ln(\dot{x}) - \ln(v_0) &= pt \\ \dot{x} &= v_0 \mathrm{e}^{pt}. \end{aligned} \tag{6.5}$$

Figure 6.2 plots eq. (6.5) with various p. Notice that when $p = 0$, eq. (6.5) is just a straight line $\dot{x} = v_0$. The more negative p is, the faster it slows down, as illustrated. That is why sometimes, we call p the **dampening factor**. Also, v_0 here just scales the graph in the $\dot{x}$ direction.

To find $x(t)$, we rewrite eq. (6.5) as

$$\frac{\mathrm{d}x}{\mathrm{d}t} = v_0 \mathrm{e}^{pt}.$$

Then,

$$\mathrm{d}x = v_0 \mathrm{e}^{pt} \, \mathrm{d}t \tag{6.6}$$

$$x + C_1 = v_0 \int \mathrm{e}^{pt}\,\mathrm{d}t. \tag{6.7}$$

Here, I shall introduce an integration technique called **change of variables**, commonly known as u-substitution. We'll formally come back to this topic later in **??**. Basically, it's a way to convert integrals that we don't recognize into an easier integral. It's better if I just show the examples. We don't know the antiderivative of e^{pt} in eq. (6.7), however we know the antiderivative of e^{u}. So let's convert e^{pt} into that form. By letting a dummy variable $u = pt - v_0$, we have to convert $\mathrm{d}u$ into $\mathrm{d}t$ aswell.

$$\begin{aligned} u &= pt - v_0 \\ \frac{\mathrm{d}u}{\mathrm{d}t} &= \frac{\mathrm{d}}{\mathrm{d}t}(pt - v_0) \\ \frac{1}{p}\,\mathrm{d}u &= \mathrm{d}t. \end{aligned}$$

Then, substituting $\mathrm{d}t = \frac{1}{p}\,\mathrm{d}u$ and $u = pt$ into eq. (6.7), it reads

$$\begin{aligned} x + C_1 &= v_0 \int \mathrm{e}^{u}\left(\frac{1}{p}\,\mathrm{d}u\right) \\ x + C_1 &= \frac{v_0}{p}\mathrm{e}^{pt}. \end{aligned} \tag{6.8}$$

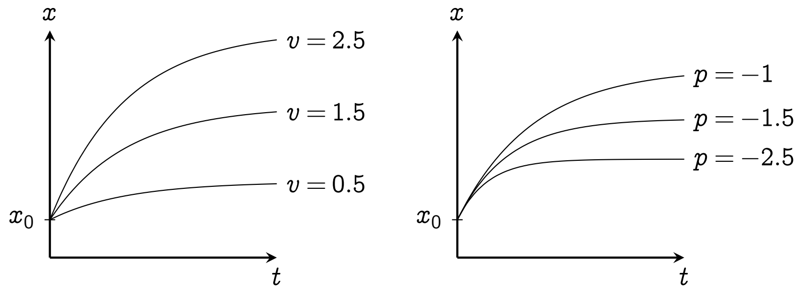
We also have to take care of the C_1. Using $t = 0 \implies x = x_0$,

$$\begin{aligned} x_0 + C_1 &= \frac{v_0}{p}\mathrm{e}^{p\cdot(0)} \\ C_1 &= \frac{v_0}{p} - x_0. \end{aligned}$$

Then, plug it into eq. (6.8). The equation becomes

$$\begin{aligned} x + \frac{v_0}{p} - x_0 &= \frac{v_0}{p}(\mathrm{e}^{pt}) \\ x &= x_0 + \frac{v_0}{p}(\mathrm{e}^{pt} - 1). \end{aligned} \tag{6.9}$$

I've plotted eq. (6.9) with varying v in fig. 6.3a and varying p in fig. 6.3b. The graph should match with what you expect intuitively: higher v will get you further, and the less the drag, the further you'll get.



(A) WITH VARYING v, SETTING $p = -1$ **(B)** WITH VARYING p, SETTING $v = 2$

FIG. 6.3 | PLOT OF EQ. (6.9), SETTING $x_0 = 0.5$.

Notice, this motion has a clear upper limit. *If* $p \in \mathbb{R}^-$, $\lim_{t\to\infty} e^{pt} = 0$. Taking the limit as $t \to 0$ on both sides of eq. (6.9), we get

$$\begin{aligned}\lim_{t\to\infty} x(t) &= \lim_{t\to\infty}\left(x_0 + \frac{v_0}{p}(e^{pt} - 1)\right).\\ &= x_0 + \frac{v_0}{p}\lim_{\cancel{t\to 0}} \cancelto{0}{\left(e^{pt}\right)} - \frac{v_0}{p} = x_0 - \frac{v_0}{p},\end{aligned}$$

which when $p \in \mathbb{R}^-$ and $v_0 \in \mathbb{R}^+$, the motion procedes forward, then gradually slows down and stops at $x_0 - \frac{v_0}{p}$ which is more than x_0.

Motion with just quadratic drag

This equation is even easier than linear drag, so I'd leave out some steps. Equation (6.3) simplifies to

$$\begin{aligned}\frac{\mathrm{d}\dot{x}}{\mathrm{d}t} &= q\dot{x}^2\\ \int \frac{1}{\dot{x}^2}\,\mathrm{d}\dot{x} &= q\int \mathrm{d}t^3\\ -\frac{1}{\dot{x}} + C &= qt.\end{aligned}$$

Taking care of C: use $t = 0 \implies \dot{x} = v_0$.

$$\begin{aligned}-\frac{1}{v_0} + C &= q \cdot 0\\ C &= \frac{1}{v_0}.\end{aligned}$$

Therefore,

$$\begin{aligned}-\frac{1}{\dot{x}} + \frac{1}{v_0} &= qt\\ \frac{\mathrm{d}t}{\mathrm{d}x} &= \frac{1 - qv_0t}{v_0}\end{aligned}$$

$$x = v_0 \int \frac{1}{1 - qv_0 t} \, dt. \tag{6.10}$$

The structure of the integral is similar to $1/t$. Therefore, let $u = 1 - qv_0 t$. Then,

$$\frac{du}{du} = du(1 - qv_0 t)$$
$$1 = -qv_0 \frac{dt}{du}$$
$$dt = -\frac{1}{qv_0} \, du.$$

Substitute $u = 1 - qv_0 t$ and $dt = -\frac{1}{qv_0} du$ into eq. (6.10):

$$x = v_0 \int \frac{1}{u} \left(-\frac{1}{qv_0} \, du \right)$$
$$= -\frac{1}{q} \int \frac{1}{u} \, du$$
$$x(t) = -\frac{1}{q} \ln(1 - qv_0 t) + C_1. \tag{6.11}$$

Taking care of C_1: use $t \to 0 \implies x(t) = x_0$.

$$x_0 = -\frac{1}{q} \ln(1 - qv_0 \cdot (0)) + C_1$$
$$x_0 = C_1$$

Plug this into eq. (6.11) to get the final answer:

$$x(t) = -\frac{1}{q} \ln(1 - qv_0 t) + x_0.$$

Surprisingly, quadratic drag does not have upper position bounds. A bit more thought would reveal that when $v < 0$, the quadratic drag $f_r(v) = a_2 v^2$ is smaller thasecn $f_r(v) = a_1 v$. Thus, it should makes sense that quadratic doesn't have bounds, but linear has an upper bound.

Terminal velocity of objects

6.1.3 Time of meteor collision from great height

Illustrated in fig. 6.4[4], a meteor is falling from height h above the Earth. Let's find the time that it'd take to hit the Earth. The meteor has the mass m, falling from height h above the ground. The Earth has mass M, radius R. Let's denote the meteor's position relative to the Earth's center with r. The initial condition of the meteor is $r(0) = h + R$, and $\dot{r}(0) = v_0$ For

[4] Not to scale.

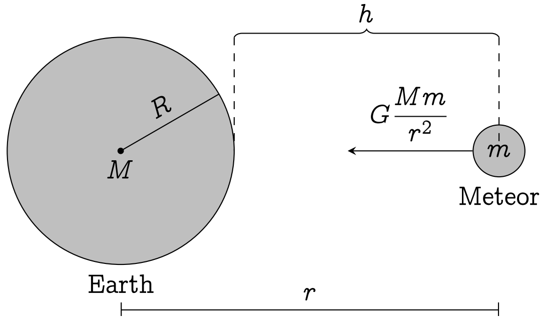


FIG. 6.4 | AN ILLUSTRATION OF A METEOR FALLING TO THE EARTH FROM HEIGHT h.

simplicity sake, let the meteor be a mass point: its radius is zero, and the Earth is very massive compared to the meteor so it doesn't move with the meteor's gravitational attraction. Given by Newton's law of gravitational attraction, the Earth is pulling the meteor by

$$F = G\frac{Mm}{r^2}.$$

Therefore, Newton's second law on the meteor reads

$$G\frac{Mm}{r^2} = m\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathrm{d}r}{\mathrm{d}t}\right)$$[5].

To frame the problem mathematically, we want to find an equation of motion of the meteor. Then, find the time it takes for the meteor to travel from $h + R$ (initial position) to R (the ground).

The m on both sides cancel. For convenience, let $\kappa = GM$

$$\frac{\kappa}{r^2} = \frac{\mathrm{d}\dot{r}}{\mathrm{d}t}. \tag{6.12}$$

Solving this equation is not at all trivial: there are *three* variables, i.e., r. $\dot{r}$, and t; however, an equation only has two sides. We can't possibly separate

[5]Notice that here, r also changes with time.

these variables. Here, I shall introduce a technique for solving this kind of differential equation. With the chain rule,

$$\frac{\mathrm{d}\dot{r}}{\mathrm{d}t} = \frac{\mathrm{d}\dot{r}}{\mathrm{d}r} \cdot \frac{\mathrm{d}r}{\mathrm{d}t} = \dot{r}\frac{\mathrm{d}\dot{r}}{\mathrm{d}r} :$$

we converted an expression that's dependent on other variable t to be dependent on a lower derivative r instead! Then t is removed, or rather, hidden. You can interpret $\frac{\mathrm{d}}{\mathrm{d}r}\dot{r}$ as the velocity at any given distance away from the Earth. Plug this into eq. (6.12), we get

$$\begin{aligned} \frac{\kappa}{r^2} &= \dot{r}\frac{\mathrm{d}\dot{r}}{\mathrm{d}r} \\ \kappa \int \frac{1}{r^2}\,\mathrm{d}r &= \int \dot{r}\,\mathrm{d}\dot{r} \\ \frac{\kappa}{r} + C &= \frac{\dot{r}^2}{2} \end{aligned} \tag{6.13}$$

The term $+C$ is going to be a different because now, we don't have a t to fix our initial condition. However, we know that when $\dot{r} = v_0$, $r = h + R$. Therefore,

$$\begin{aligned} -\frac{\kappa}{h+R} + C &= \frac{v_0^2}{2} \\ C &= \frac{v_0^2}{2} - \frac{\kappa}{r}. \end{aligned}$$

However, the structure of C is quite complicated so, I wouldn't substitute it in yet until we get our final answer. Continuing with eq. (6.13), we turn the $\dot{r}$ into the Leibniz's notation:

$$\begin{aligned} \frac{\kappa + Cr}{r} &= \frac{1}{2}\left(\frac{\mathrm{d}r}{\mathrm{d}t}\right)^2. \\ \sqrt{2}\sqrt{\frac{\kappa + Cr}{r}} &= \frac{\mathrm{d}r}{\mathrm{d}t} \\ \int \mathrm{d}t &= \sqrt{2}\int \sqrt{\frac{r}{\kappa + Cr}}\,\mathrm{d}r \end{aligned}$$

Unfortunately, this integral is very hard to solve. But it is possible, and the solution to this integral is

$$\frac{C}{\kappa^{3/2}\sqrt{r}}\sqrt{\frac{r}{\kappa r + C}}\sqrt{\frac{\kappa r}{C}+1}\left(\sqrt{\kappa r}\sqrt{\frac{\kappa r}{C}+1}-\sqrt{C}\sinh^{-1}\left(\sqrt{\frac{\kappa x}{C}}\right)\right),$$

which is quite a nightmare, but we will get back to this in the far far future.

6.1.4 Damped harmonic motion

6.2 Conservation laws

Part II

Signal analysis

Part III

The extensions

Part IV

The fundamentals, reimagined: real analysis

CHAPTER 7

Constructing the real numbers

Abstract

Prerequisites: *intuitions of set theory, basic discrete math*

As far as the "calculus" part of this book goes, it doesn't really delve deep into the proof: the backbones of how structures work together. For example, how do you know that in the definition of derivatives, if $h \to 0$, the value actually converges to something. To be frank, real analysis is quite abstracted away from the physical reality therefore, it's a bit dry. This is quite the double edged sword of math. With real analysis, we have the power to tell clearly if something is true or not. However, most times it's mistakenly used to the bones: too abstract that the learner does not have any clear concepts leftover. All the rest is just some meaningless mathematical notation that's floating in the air. And I don't want that.

The goal for the real analysis part of this book is to provide an enjoyable experience delving in to the proofs behind the backbones of calculus. Therefore, I shall try to illustrate everything with diagrams so it's simple to visualize and not too abstracted away from reality. Now that you know my intentions, let's start.

7.1 The mindset of real analysis

Before we study the reals, we must know the mindset of real analysis first. Analysis is used to generalize and study the *exact* behaviors of mathematical entities. In real analysis, we study the *reals*. Most of the stuffs in mathematics were built way before real analysis. However, it's not rigorous and it's prone to error. Here, real analysis comes to play.

We *abstract* properties of mathematical identities away from the numbers, and we generalize it. But we can't just choose everything, we must be very wise. The properties that we select to be true are called **axioms** After all the decision has been done, we must find the most general mathematical entity that satisfies it. And thus, we shall begin with the most basics of analysis: set theories.

7.2 The Zermelo–Fraenkel set theory

In here, we shall explore what's the backbones of sets that will lead to the mechanics of numbers. And here arises the set theory. Firstly, a **set** is a group of things, whether it be mathematical entities or real world objects. If two sets contains the same elements, then it's the same set. That means, set does not care about permutation. A wiser way to state this is

Axiom 1: Axiom of Extensionality

Two sets are the same if they have the same elements.

$$\forall X \forall Y [\forall z (z \in X \iff z \in X) \implies X = Y]. \tag{7.1}$$

Translation: Set X and Y will be equal iff for all elements z, z is in both X and Y.

which just means that "A *set is uniquely determined by its members*".

Then, we also have to define that a set cannot have the same elements that is,

Axiom 2: Axiom of foundation

Every non-empty set x contains a member y such that x and y are *disjoint*.

$$\forall x [x \neq \emptyset \implies \exists y ((y \in x) \wedge (y \cup x) = \emptyset)] \tag{7.2}$$

Translation: For all non-empty set x, there exists y where both y

Part V

Beyond imagination: complex analysis

APPENDIX A

Fundamental of physics

APPENDIX B

The binomial theorem

Index

Bibliography

[1] M. Alkousa, *Mit integration bee, solutions of qualifying tests from 2010 to 2023, 249 pages, 15 chapters.* (Independently Published, May 22, 2023).

[2] J. Foster, *U-substitution*, (visited on 09/26/2024).

[3] Math Visualized, *Galois theory explained simply*, Nov. 10, 2020.

[4] J. Turnbull, D. Lea, D. Parkinson, P. Phillips, B. Francis, S. Webb, V. Bull, and M. Ashby, *Oxford advanced learner's dictionary, 8th edition: paperback*, Oxford Advanced Learner's Dictionary, 8th Edition (OUP Oxford, 2010).

[5] Wikipedia contributors, *Summation: powers and logarithm of arithmetic progressions*, (Feb. 4, 2024) https://en.wikipedia.org/wiki/Summation#Powers_and_logarithm_of_arithmetic_progressions.