Compact Calculus

with relation to existence

Puripat Thumbanthu

1E

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

Contents

Preface

Acknowledgement

Reading Guide

Big disclaimer. I am a physicist, and this is calculus from the physicist's point of view. So, there's going to be some contents from physics that pure mathematicians might not adore, I'm terribly sorry for that.

The abstract contains the guide, the overview, and the mindset of the chapter. *Reading the abstract is a necessity.*

The chapter contains the main idea of the topic. A chapter might have supplementary unnumbered chapters or sections which is optional.

Appendices either supplements or extend the chapter. Some even goes much beyond the chapter. They all vary. It's recommended to study them separately.

The interludes provide a *historical background* for the next chapter. It's meant to connect one chapters with the other. It acts as a storyline bridge. If not taken, the next chapter might seem too terse. So, I recommend the reader to skim through the fruitfulness of historical development.

This book is separated into five parts

I The fundamentals

II The applications

III The extensions

Part I (The fundamentals) focuses on the basics of calculus: from derivatives to antiderivatives and some of its applications. I've swapped around the order of contents a lot as I see fit. I've also introduced some applications that's not found of normal pedagogy.

Part II (The applications) focuses on further applications calculus to real world problems, mostly in physics.

Part III. (The extensions) explores the realm of specialized calculus, most of the aren't even taught in universities: newer branches of calculus.

Part IV. (Real analysis) and *Part V. (Complex analysis)* as the name suggests, explores the full behaviors of real and complex functions. It reconsiders all the basics of calculus and dives deep into the backbone of all symbols that are abstracted away from the physical world.

At last, a fair warning; not all contents follow accurate historical order.

Part I

The foundations

CHAPTER 1

Writing down nature: derivatives

1.1 Invitation to calculus from a physicist

Prerequisites: *graphs, functions, kinematic variables*

1.1.1 The mission of calculus

Calculus is in two parts: differential and integral, which is linked by the fundamental theorem of calculus. Calculus *links* physical systems to mathematical equations, so we can *predict* the future of a system from the solution of those equations. But before we *predict*, we must describe the system first: we must find a *link*. And that's what we're going to do in this chapter.

The system that we'll use in this chapter is a mass that's traveling with some speed. We'll see how the attempt to describe this system naturally give birth to *derivatives*. In the next chapter, we'll find the solution to the equation we've written. And at the end, we'll develop an intuition that

1. Derivatives 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

1.1.2 The notation of derivative 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 $\mathbf{v}$.

1.1.3 Message from a physicist about the pedagogy

Physics is about trying to understand how nature works. What we do in physics is that we first describe a system using mathematical equations, then use the solution of that equation to reveal the nature of that system. Here, I'd do the same.

If I were teaching a class, I'd tell you about car design. But, this is a textbook, and trying to deliver the same story here isn't that concise and cohesive. So, instead of following a story about cars, I'd like you to hop on with me and analyze the mechanics of a **speedometer**, which I reckon is the best introduction to calculus there is.

1.2 Speed and instantaneous rate of change

Let's analyze how a car **speedometer** 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 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

FIG. 1.1 | CALCULATING THE VELOCITY OF A CAR FROM MOTION BLUR.

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 1.1.2, 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"

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

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

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

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

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 $\frac{f(x+h)-f(x)}{h}$ 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

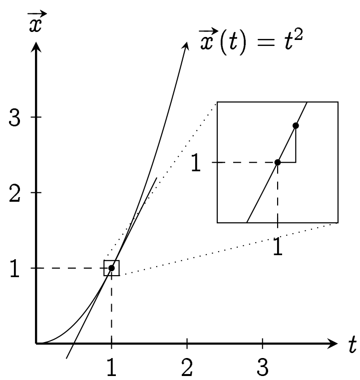


FIG. 1.2 | EXAMPLE OF POSITION VS. TIME GRAPH WHERE $\mathbf{x}(t) = t^2$.

To interpret derivatives, notice that Equation (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?

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

Here, I'll give you an example of evaluating the derivative by using the definition, a.k.a. the **method of increments.**

In fig. 1.2, the function $\mathbf{x}(t)$ is just t^2: the position at any time t is t^2. I'd evaluate the velocity at $t = 1\,\mathrm{s}$. 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

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

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.

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 ball, and $\mathbf{g}$, the acceleration from Earth's gravity. Newton's second law tells us that

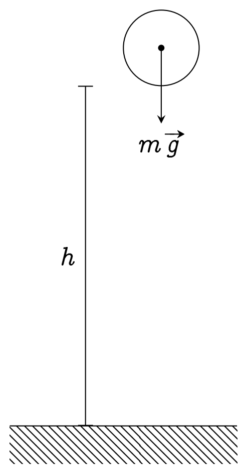


FIG. 1.3 | A BALL DROPPED FROM HEIGHT h

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

$$m\mathbf{g} = m\frac{\mathrm{d}^2\mathbf{r}(t)}{\mathrm{d}t^2} \quad \text{or,} \quad \mathbf{g} = \frac{\mathrm{d}^2\mathbf{r}(t)}{\mathrm{d}t^2}, \tag{1.12}$$

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*

Before we dive in, I shall clarify that every line, including straight, is a mathematical **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* [1], 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: mission impossible

2.1.1 The mindset of integral calculus

In the previous chapter, we've learned how to describe the universe using derivatives. But derivatives falls short when we want to predict the future of a system. But, what do we mean when we say "predict the system?"

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 RHS 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

In eq. (2.1), both $\mathbf{r}$ and t is mixed up on the same side of the equation. That's not good for solving equations. So let's separate them. First, 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.

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

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

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.

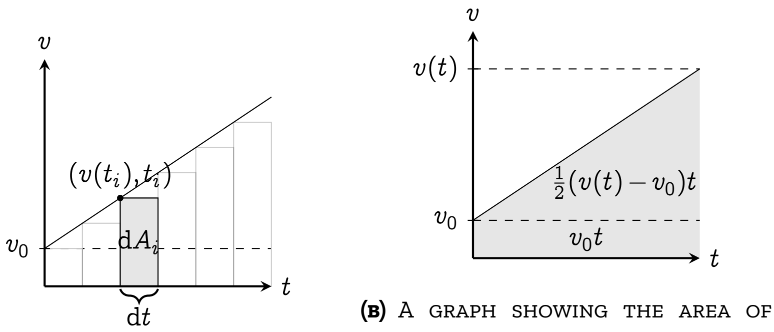
$$\mathbf{v} = \frac{\mathrm{d}\mathbf{r}}{\mathrm{d}t} = \mathbf{g}t + \mathbf{v}_0 \tag{2.13}$$

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

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, $\mathrm{d}t\,(\mathbf{g}t + \mathbf{v}_0)$ would just represent an area of a little rectangle $\mathrm{d}A$ as shown in the figure.

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

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

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

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

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,

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

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

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

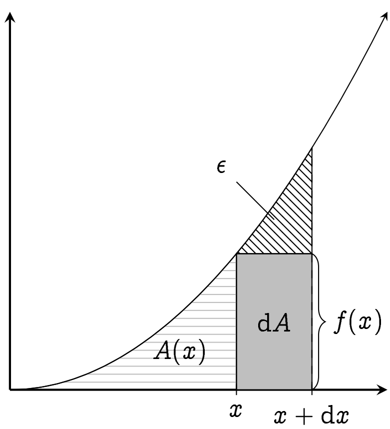


FIG. 2.2 | THE GEOMETRICAL INTERPRETATION OF PART ONE OF THE FUNDAMENTAL THEOREM OF CALCULUS (THEOREM 1).

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

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

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

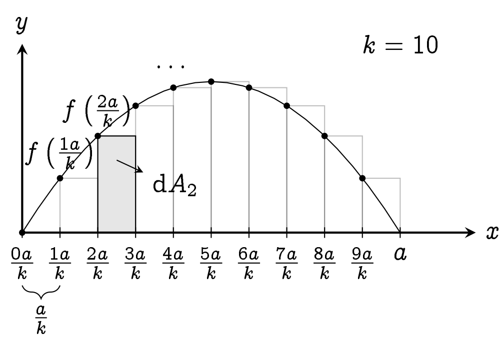


FIG. 2.3 | ILLUSTRATION OF RIEMANN SUM OF A FUNCTION $f(x)$ FROM 0 TO a BY SETTING $k = 10$

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

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

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* [2].

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

$$\sum_{i=0}^{k} \frac{a}{k} \times f\left(\frac{ia}{k}\right) = \sum_{i=0}^{k} \frac{a}{k} \times 4\left(\frac{ia}{k}\right)^3$$

$$= 4\sum_{i=0}^{k}\left(\frac{x}{k}\right)^4 i^3$$
$$= 4\left(\frac{x}{k}\right)^4\left(\frac{k(k+1)}{2}\right)^2.$$

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

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

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 move at acceleration 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

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

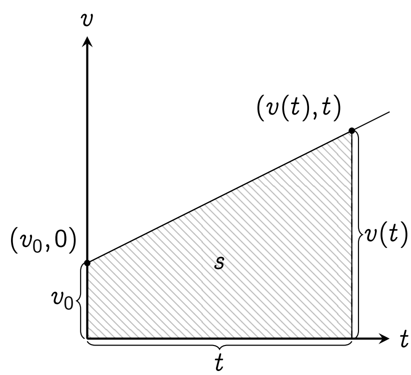


FIG. 2.4 | v-t GRAPH OF AN OBJECT UNDER CONSTANT OR UNIFORMED ACCELERATION

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

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

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

So, we dissect a circle radially into small rings dr thin, as shown 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 lit-

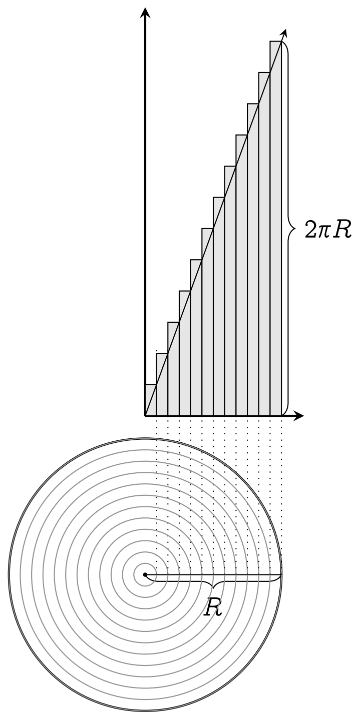


FIG. 2.6 | REARRANGING ALL THE APPROXIMATED RECTANGLES ONTO A GRAPH (NOT TO SCALE).

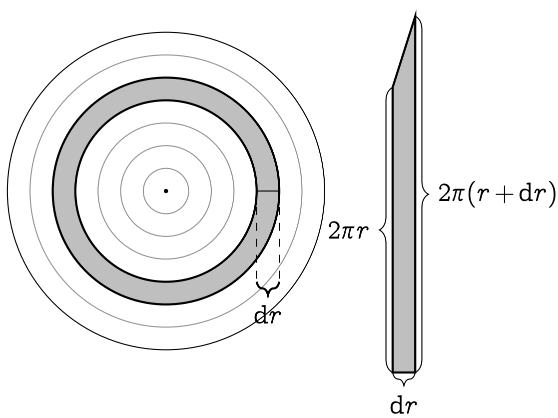


FIG. 2.5 | (LEFT) DISSECTING A CIRCLE RADIUS R RADIALLY INTO RINGS, EACH RING dr THIN. (RIGHT) STRETCHING A RING INTO A TRAPEZOID (NOT TO SCALE).

tle trapezoid dA 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 d$r \to 0$, dr^2 becomes negligible. Therefore,

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

This equation says that for d$r \to 0$, the trapezoid becomes a rectangle side length dr 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].

For every stripe at r, its height is $2\pi r$. If you were to plot the height of all rectangles when 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

Abstract

This chapter focuses on evaluating the derivative and integrals of common functions such as polynomials and exponentials. *The geometrical interpretation of both derivatives and integrals plays a massive role in this chapter.* I shall warn you a bit, we have rules for differentiation, but not for integration. Integration is an art of mathematical manipulation, and is discussed further in chapter 7 and chapter 8.

We'll go through the derivatives and antiderivatives of

1. Monomials (ax^n) and Polynomials ($a_0 + a_1x^1 + a_2x^2 + \dots$)
2. Exponential functions (an^x)
3. Logarithmic functions ($\log_n(ax)$)

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

Note on terminology The integral refers to the area under the curve. The antiderivative refers to the function that outputs the area under the graph. We *evaluate* the integral, but we *find* the antiderivative of a function.

3.1 Trivial rules

3.1.1 The chain rule

The Leibniz' notation treats derivative as fractions. You can cancel terms as seen earlier in eq. (1.9). This property can be used to take derivatives of composite functions. E.g., finding the derivative of $f(g(x))$ but you only know $f(x)$ and $g(x)$. First, substitute $g(x)$ as u.

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

We know that *one* is the multiplicative identity, and *one* is any number divided by itself[1]. Let $1 = \frac{\mathrm{d}u}{\mathrm{d}u}$, then

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

Performing change of denominator, then substitute $u = g(x)$:

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

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

[1] Except zero, of course.

is 4 times the walking speed. The ratio between the cheetah's speed compared to walking speed would obviously be $10 \cdot 4 = 40$:

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

3.1.2 Integral constant

In **??**, each time that we reverse the derivative, i.e., find the antiderivative, we add an initial condition term, i.e., v_0 and r_0. This is not optional. You have to do this every time you want to find the antiderivative of something. If $A(x)$ is the antiderivative of $f(x)$, that means

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

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

3.1.3 Integrals of the infinitesimal

As seen in definition 2, integrating a small area gives you the whole area. The antiderivative of the small rectangles is just the whole area plus the integral constant.

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

3.2 Trivial rules

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 depend on any value, therefore the derivative of a constant c is zero:

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

3.3 Linearity of differentiation and integration

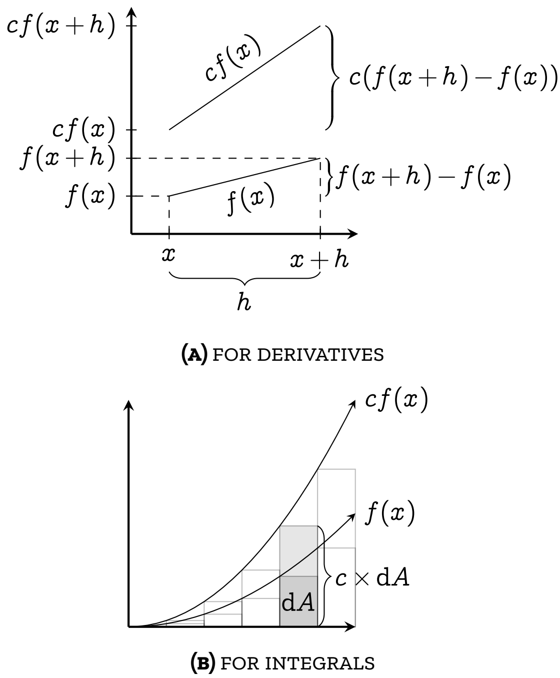


FIG. 3.1 | CONSTANT MULTIPLE RULES

It shouldn't be too hard to think of these rules visually because it's just scaling and adding functions together. For any function $f(x)$ and $g(x)$, and constants a and b, the following rules follows.

The constant multiple rules:

$$\frac{\mathrm{d}}{\mathrm{d}x}(af(x)) = a\frac{\mathrm{d}}{\mathrm{d}x}f(x), \qquad \text{Illustrated in fig. 3.1a} \tag{3.6}$$

$$\int af(x)\,\mathrm{d}x = a\int f(x)\,\mathrm{d}x\,. \qquad \text{Illustrated in fig. 3.1b} \tag{3.7}$$

The sum rules:

$$\frac{\mathrm{d}}{\mathrm{d}x}(f(x)+g(x)) = \frac{\mathrm{d}}{\mathrm{d}x}f(x) + \frac{\mathrm{d}}{\mathrm{d}x}g(x), \tag{3.8}$$

$$\int f(x)+g(x)\,\mathrm{d}x = \int f(x)\,\mathrm{d}x + \int g(x)\,\mathrm{d}x\,. \tag{3.9}$$

The slope of the sum of two functions adds up, and so the area.

Generally, we say that these operations, i.e., the derivative and the antiderivative is **linear** because they have the property

$$\mathrm{D}(af(x)+bg(x)) = a\mathrm{D}(f(x)) + b\mathrm{D}(g(x)),$$

where D is an **operator**. An operator is like an instruction to do something. Here, D could represent "*take the derivative*", or "*find the antiderivative*". You could interpret "*linear*" as in derivatives approximate everything as a line (Remarks 1.1). The true meaning of this is actually engrained in linear algebra which is beyond the scope of this book. You can consult any linear algebra textbook for it.

3.4 Derivatives and antiderivatives of polynomials

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

$$f(x) = a_0 + a_1x + a_2x^2 + a_3x^3 + \dots\,.$$

Because the linearity of both derivatives and antiderivatives, we can break down the polynomial into multiple monomials. Thus, if we know the derivative and antiderivatives of x^n, we basically get the whole family for free.

3.4.1 Derivatives of polynomials: the power rule

Let's start of with the derivatives. We could use the method of increments every time, but it isn't quite interesting. Therefore, let's focus on the geometrical intuition first by considering the derivative of x^2[2]. Geometrically, x^2 resembles a square. The derivative question then becomes "Take a square side length x then, increase its side length by $\mathrm{d}x$. How much area has changed in proportion to $\mathrm{d}x$."

Illustrated in fig. 3.2, we can say that

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(x^2\right) = \lim_{\mathrm{d}x \to 0} \frac{x\,\mathrm{d}x + x\,\mathrm{d}x + \mathrm{d}x^2}{\mathrm{d}x}. \tag{3.10}$$

Geometrically, the $\mathrm{d}x^2$ square is negligible. When $\mathrm{d}x \to 0$, that square will just become a single point compared to the big $x\,\mathrm{d}x$ on the side; therefore, we can safely leave it out. The equation then becomes

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

the derivative of x^2 is $2x$.

[2] Don't worry, if we can find its derivative, we can also find its antiderivative after.

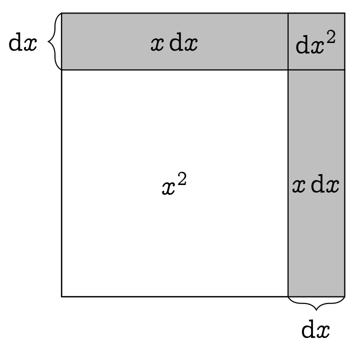


FIG. 3.2 | THE GEOMETRICAL INTERPRETATION OF THE DERIVATIVE OF x^2

Notice, eq. (3.10) has the same form as if we were to use the method of increments

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

This shows that the geometrical method and the symbolical method is equivalent in nature. Now, deriving the derivative of x^3 geometrically shouldn't be hard either. Just take a cube side length x, then increase its side length by $\mathrm{d}x$ and see how the volume changes in proportion to $\mathrm{d}x$. The final answer should be $3x^2$.

We could go on and derive the derivative of higher powers geometrically using 4-dimensional space. But sadly, we don't have the glory of higher dimensions in our world[3]. To go further without headaches, we turn to the binomial expansion and the mighty definition of derivatives

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

Then, use the binomial expansion (appendix B) on $(x+h)^n$.

$$\begin{aligned}
&= \lim_{h\to 0}\frac{\left(\sum_{k=0}^{n}\binom{n}{k}x^{n-k}\cdot h^n\right) - x^n}{h} \\
&= \lim_{h\to 0}\frac{\binom{n}{0}x^n h^0 + \binom{n}{1}x^{n-1}h^1 + \binom{n}{2}x^{n-2}h^2 + \ldots \binom{n}{n} + x^0 h^n - x^n}{h} \\
&= \lim_{h\to 0}\frac{x^n + nx^{n-1}h^1 + \binom{n}{2}x^{n-2}h^2 + \cdots + h^n - x^n}{h} \\
&= \lim_{h\to 0} nx^{n-1} + \binom{n}{2}x^{n-2}h + \binom{n}{3}x^{n-3}h^2 + h^{n-1}
\end{aligned}$$

As $h \to 0$, the remaining terms would vanish. Thus,

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(x^n\right) = nx^{n-1}, \tag{3.11}$$

[3]Technically, it's possibly to visualize 4-dimensional space. However, it'd take months to get the hang of it. Not even mentioning the fifth dimension. Therefore, we should leave it off for now.

which is what we call the power rule of derivative. I'd like to clarify that binomial theorem does work for all real numbers. The prove is in appendix B if you'd like to take a look for yourself.

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

1. $\frac{\mathrm{d}}{\mathrm{d}x}\left(x^2 - 2x + 16\right)$ $\qquad 2x - 2$

2. $\frac{\mathrm{d}}{\mathrm{d}x}\left(x^3 + x^2 + x + 1\right)$ $\qquad 3x^2 + 2x + 1$

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

3.4.2 Antiderivatives of polynomials: the reversed power rule

Now, we can move on to reversing the power rule. We want to find the antiderivative of x^n. By the fundamental theorem of calculus and eq. (3.11),

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

Since n is a dummy variable, we can change n to $n+1$

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

and by the linearity of integrations (Equation (3.7)), we get the **reversed power rule**:

$$\int x^n \,\mathrm{d}x = \frac{x^{n+1}}{n+1}. \tag{3.12}$$

Notice, this rule does not work for $n = 1$ because we can't divide by zero...*or can we???* (Discussed further in section 3.9)

3.5 Extending the equations of linear motion

In section 2.5.1, we discussed the equation of motion of objects with uniformed acceleration. What if now, the acceleration is changing over time? In a sense, we can use the Newton's second law to deal with that also. If the jerk j is constant, the acceleration must be increasing uniformly

$$j = \frac{\mathrm{d}a}{\mathrm{d}t}.$$

We can then move $\mathrm{d}t$ around and integrate both sides by using the reversed power rule on it.

$$\int j \,\mathrm{d}t = \int \mathrm{d}a$$

$$j \int \mathrm{d}t = \int \mathrm{d}a \qquad \text{Linearity of integrals}$$

$$jt + a_0 = a \qquad \text{Integrating an infinitesimal \& Integral constant}$$

Then, because a is the derivative of v. Note that both a, and v is a function of time.

$$\frac{\mathrm{d}v}{\mathrm{d}t} = jt + a_0$$

$$\int \mathrm{d}v = \int jt + a_0 \,\mathrm{d}t$$

$$v = j \int t \,\mathrm{d}t + \int a_0 \,\mathrm{d}t \qquad \text{Linearity of integrals}$$

$$v = \frac{1}{2} jt^2 + a_0 t + v_0. \qquad \text{Reversed power rule \& Integral constant}$$

Because v is the derivative of r,

$$\begin{aligned}
\frac{\mathrm{d}r}{\mathrm{d}t} &= \frac{1}{2} jt^2 + a_0 t + v_0 \\
\int \mathrm{d}r &= \int \frac{1}{2} jt^2 + a_0 t + v_0 \,\mathrm{d}t \\
r &= \frac{1}{2} j \int t^2 \,\mathrm{d}t + a_0 \int t \,\mathrm{d}t + v_0 \int \mathrm{d}t \\
r &= \frac{1}{6} jt^3 + \frac{1}{2} a_0 t^2 + v_0 t + r_0.
\end{aligned}$$

And there you have it! Technically we can extend this to whatever we want. The jerk also doesn't have to be uniformed, it could be a function of time itself. But, that's probably enough. If you'd like to try, extend the equation of motion to uniformed snap s. The final equation should be

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

3.6 Exponentials, growth pill

Let's say there's a magical drop of water that doubles its volume V every hour. That means, for every time t,

$$V(t+1) = 2V(t). \tag{3.13}$$

We can find the function $V(t)$ that satisfies the function above. Because we're dealing with integers time here, we can consider the function from $V(0)$ to $V(t)$. If we set $V(0) = 1$, that is the drop starts at one unit of volume,

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

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	$1024 - 512 = 512$

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

It's clear that the pattern is $V(t) = 2^t$, which is an exponential function. Because we're in calculus, a natural question to ask is "*what is its rate of*

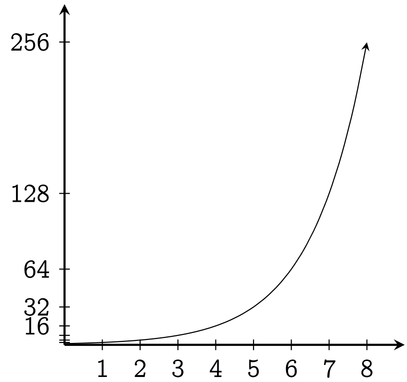


FIG. 3.3 | EXPONENTIAL FUNCTION 2^x PLOTTED FROM 0 TO 8

change?". So you might start plotting it over time (fig. 3.3). However, these exponential grows so quickly that by $V(7)$, we're already in the hundreds. It'd probably be better to list the values of each point on a table. On the right most column of table 3.1, there's something suspicious going on. Specifically, the difference between $V(t)$ and $V(t-1)$ is exactly $V(t-1)$: the function changes as much as its past-self. Does that mean that the derivative of 2^x equals 2^x?

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

$$\frac{2^{x+1}-2^x}{1} = 2^x\left(\frac{2-1}{1}\right) = 2^x, \tag{3.14}$$

that is why we saw that $V(t)-V(t-1) = V(t-1)$. However, if we'd want to calculate the derivative of 2^x, you'd have to use the method of increments:

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

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.

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(3^x\right) = 3^x \lim_{h\to 0} \frac{3^h - 1}{h}. \tag{3.15}$$

The rate of change of an exponential function is always itself times a proportionality constant. For 3^x, it's about 1.09851 … . 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!

3.6.1 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[4] called the **power series**. For every $f(x)$,

$$f(x) = a_0 + a_1 x^1 + a_2 x^2 + a_3 x^3 + \dots . \tag{3.16}$$

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

$$\sin(x) = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \frac{1}{7!}x^7 + \dots . \tag{3.17}$$

we will derive this expression later in section 4.4. For now, we can just use eq. (3.16) 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,

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(n^x\right) = n^x. \tag{3.18}$$

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

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

[4] 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.

$$n^x = a_0 + a_1 x^1 + a_2 x^2 + a_3 x^3 + \dots, \tag{3.19}$$

$$n^x = a_1 + 2a_2 x^1 + 3a_3 x^2 + 4a_4 x^3 + \dots. \tag{3.20}$$

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

$$\begin{aligned} a_0 &= a_1 \\ a_1 &= 2a_2 \\ a_2 &= 3a_3 \\ a_3 &= 4a_4 \\ &\vdots \end{aligned} \tag{3.21}$$

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. (3.20):

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

Since $a_0 = a_1$, a_1 must also be 1. We can then go back to eq. (3.21) 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 $2.71828\ldots$ has a name and symbol: the Euler's number[5], written as e where

$$\mathrm{e} = \sum_{i=0}^{\infty} \frac{1}{i!} \tag{3.22}$$

3.6.2 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

$$n(t) = n_0 + tr \tag{3.23}$$

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:

$$n(t+1) = n(t)r + n(t). \tag{3.24}$$

We can find the expression for $n(t)$ in a similar fashion to what we've done in eq. (3.13). You'll get

$$n(t) = (1+r)^t n_0 \tag{3.25}$$

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

[5] Not to be confused with the "Euler's constant" which is another constant written γ, and is around $0.57721\ldots$

option 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. (3.25), 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

$$x = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n \tag{3.26}$$

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. (3.26) equals eq. (3.21), we'd find the upper limit for this problem and solve the mystery.

We can use the binomial theorem on eq. (3.26) 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.

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

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

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

$$x = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \ldots \tag{3.27}$$

which is literally eq. (3.21). That means, the upper limit that the bank can give you is e. Geometrically, it should make sense. Because we're gradually turning a discrete interest into a continuous one, e should appear in the limit of the continuous bank interests.

3.6.3 The antiderivative of exponential functions

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

$$\int \mathrm{e}^x \,\mathrm{d}x = \mathrm{e}^x + C. \tag{3.28}$$

For other bases, we could use theorem 1, the fundamental theorem of calculus, to find its antiderivative. That is, if

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

then

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

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

3.7 Logarithms

Monomials has their inverse functions: the roots, exponentials also has an inverse functions: the logarithms. Here's a simple example to illustrate what I mean.

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

Logarithms are inverses functions of exponentials: it cancels exponentials. With it comes the following properties:

$$\log_a(x) + \log_a(y) = \log_a(xy) \tag{3.29}$$

$$\log_a(x) - \log_a(y) = \log_a\left(\frac{x}{y}\right) \tag{3.30}$$

$$\log_a(x) = \frac{\log_b(x)}{\log_b(a)} \tag{3.31}$$

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 $\ln(x)$.

So what's the growth of $\ln(x)$? At first sight, since the exponential grows so fast, the inverse of exponentials must grow very slowly. It might just be the reciprocal of x: $1/x$. However, we have to derive it somehow. You could try the method of increments and get

$$\frac{\mathrm{d}}{\mathrm{d}x}(\ln(x)) = \lim_{h\to 0}\frac{\ln(x+h) - \ln(x)}{h}. \tag{3.32}$$

As seen, there are no properties of logarithm that can manipulate this. However, I shall introduce quite a sneaky concept here called **implicit differentiation**.

Notice that we've only concerned the function that's clearly written in the form $f(x) = y$ and its rate of change. We call these neat functions **explicit functions**. However, not all functions are written in this form. E.g., $\sqrt{x^2 + y^2} = 2$. These functions are called **implicit functions**, and we can actually differentiate it.

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

$$\mathrm{e}^{\ln(x)} = \mathrm{e}^{y}. \tag{3.33}$$

Because logarithms are inverses of exponentials,

$$x = \mathrm{e}^{y}. \tag{3.34}$$

If you remember from section 3.2, if two sides of the equation are equal, their derivatives with respect to the same variable must also be equal. We can take its derivative with respect to y instead of x:

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

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 in the y:

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

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.

$$n^x = \left(\mathrm{e}^{\ln(n)}\right)^x = \mathrm{e}^{x\ln(x)}. \tag{3.35}$$

With the chain rule, discussed in section 3.1.1, we can let $u = x\ln(n)$ and add an intermediate step:

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

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

$$\ln(n) = \lim_{h\to 0} \frac{n^h - 1}{h}. \tag{3.36}$$

The antiderivative of other bases exponents are then given by

$$\int n^x = \frac{1}{\ln(n)} n^x + C. \tag{3.37}$$

3.7.1 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.

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,

$$\frac{\mathrm{d}}{\mathrm{d}x}(f(x)g(x)) = \lim_{h\to 0} \frac{g(x)\Delta f + f(x)\Delta g}{h}$$

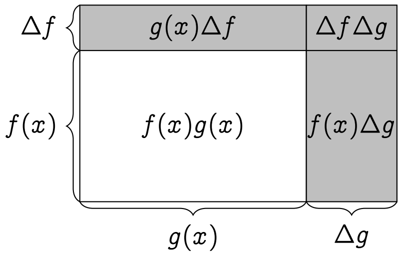


FIG. 3.4 | THE GEOMETRICAL INTERPRETATION OF THE PRODUCT RULE.

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

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

$$
\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{f(x)}{g(x)}\right) = \frac{f(x)\frac{\mathrm{d}g(x)}{\mathrm{d}x} - g(x)\frac{\mathrm{d}f(x)}{\mathrm{d}x}}{f(x)^2}. \tag{3.38}
$$

But how's about the product of three functions, e.g., $a(x)b(x)c(x)$? Or even more functions multiplied together? Well we ran into the same problem as what we were doing earlier in section 3.4.1, deriving the power rule: we don't have enough dimensions. So sadly, we have to turn ourselves to analytical method.

Let's think of this through. We still don't know how to take derivatives of products of multiple functions. However, we know the derivatives of *sums* of multiple functions by the linearity property. So if we can turn products into sum, this problem would be our candy. Gladly, there's a function that can do exactly that: the logarithms. 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.

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

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

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

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

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

And there we have it: the generalized product rule.

3.7.2 Alternate derivations for the power rule

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

$n\ln(x)$.

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

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

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

3.8 Implicit differentiation

We've discussed that we can differentiate implicit functions. Normally, it wouldn't be quite useful, but this **implicit differentiation** shows up when two rates of change are related to each other: the **related rates** problem.

Take the example of a sliding ladder, sketched in fig. 3.5. The ladder is 5m long. If any one moment, what is the rate of sliding along the x-axis w.r.t. the y-axis?

The problem is asking us to find $\frac{\mathrm{d}x}{\mathrm{d}y}$. Here, the rates 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 clearly related because the ladder length still stays the same over time. If y decreases, x must increase. Both variables are related by the pythagorean theorem

$$x^2 + y^2 = 5^2. \tag{3.39}$$

Then, we can differentiate this w.r.t. t, and use the chain rule

$$\frac{\mathrm{d}x^2}{\mathrm{d}t} + \frac{\mathrm{d}y^2}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}(5^2)$$

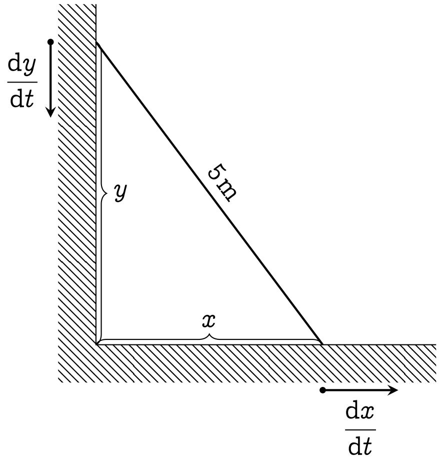


FIG. 3.5 | A LADDER LENGTH 5m SLIDING DOWN A CORNER.

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

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.

3.9 But why is the integral of the reciprocal the natural logarithm?

As we've seen, the antiderivative of $\frac{1}{x}$ is $\ln(x)$ by the fundamental theorem of calculus (theorem 1). I, however, find it disturbing and unresolved. It's a hole in the reversed power rule. From this dissatisfaction, I spent a night coming up with a way to derive this using just the reversed power rule. Enjoy the transformation!

Function	Antiderivative
x^{-3}	$-x^{-2}/2$
x^{-2}	x^{-1}
x^{-1} or $1/x$	$\ln(x)$
x^0 or 1	x
x^1	$x^2/2$

FIG. 3.6 | TABLES OF REVERSED POWER RULE FROM x^{-3} TO x^1

$$
\begin{aligned}
\int \frac{1}{x}\,\mathrm{d}x &= \int \lim_{h\to 0}\left(\frac{1}{2}x^{-1+h} + \frac{1}{2}x^{-1-h}\right) \\
&= \lim_{h\to 0}\int\left(\frac{1}{2}x^{-1+h} + \frac{1}{2}x^{-1-h}\right) \\
&= \lim_{h\to 0}\int\left(\frac{1}{2}\frac{x^{-1+h+1}}{(-1+h+1)} + \frac{1}{2}\frac{x^{-1-h+1}}{(-1-h+1)}\right) \\
&= \lim_{h\to 0}\left(\frac{1}{2}\frac{x^h}{h} - \frac{1}{2}\frac{x^{-h}}{-h}\right) = \lim_{h\to 0}\left(\frac{1}{2}\frac{x^h}{h}\frac{x^h}{h} - \frac{1}{2}\frac{1}{hx^h}\right) \\
&= \lim_{h\to 0}\left(\frac{x^{2h}-1}{2hx^h}\right) = \lim_{h\to 0}\left(\frac{\mathrm{e}^{2h\ln(x)}-1}{2h\ln(x)}\cdot\frac{\ln(x)}{x^h}\right) \\
&= \lim_{h\to 0}\left(\frac{e^{2h\ln(x)}}{2h\ln(x)}\right)\cdot\lim_{h\to 0}\left(\frac{\ln(x)}{x^h}\right) \\
&= \lim_{h\to 0}\left(\frac{e^{2h\ln(x)}}{2h\ln(x)}\right)\cdot\ln(x).
\end{aligned}
\tag{3.40}
$$

Then, we evaluate the limit at the front by letting $u = 2h\ln(x)$. When $h \to 0$, $u \to 0$ aswell. Then, use the definition of e from eq. (3.26).

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

Change the limits from $n \to \infty$ into $n \to 0$. Notice, $\lim_{n\to 0}\left(1+\frac{1}{n}\right)^n = \lim_{n\to 0}\left(1+n\right)^{1/n}$. 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. (3.40), you'll see that

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

3.10 Formula for Chapter 3

3.10.1 Formula for derivatives of functions

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

3.10.2 Formula for antiderivatives of functions

1. $f(x) = g(x) \implies \int f(x)\,\mathrm{d}x = \int g(x)\,\mathrm{d}x$ (Rules of Equality)
2. $\int af(x) + bg(x)\,\mathrm{d}x = a\int f(x)\,\mathrm{d}x + b\int g(x)\,\mathrm{d}x$ (Linearity of integration)
3. $n \neq -1$, $\int ax^n\,\mathrm{d}x = a\frac{x^{n+1}}{n+1} + C$ (Reversed power rule)
4. $\int n^x\,\mathrm{d}x = \frac{1}{\ln(n)}n^x + C$, $\int \mathrm{e}^x\,\mathrm{d}x = \mathrm{e}^x + C$ (Antiderative of exponentials)
5. $\int \frac{1}{x}\,\mathrm{d}x = \ln(x) + C$ (Antiderivative of natural logarithms)

3.10.3 Definition for various functions and constants

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

CHAPTER 4

Basic calculus and trigonometry

Abstract

Here, I assume you already have a knowledge of trigonometry. This chapter is built to reinforce that and unify trigonometry with calculus. I still try to keep the same theme throughout; having the geometrical interpretation vivid. This chapter will explore

1. The physical interpretation of derivatives and integrals of trigonometric function
2. The relationship between trignometric functions and exponentials
3. Newton's law in other coordinates

We'll extensively use trigonometry in chapter 5. Also, this chapter might seem a bit dry because there are no examples that I can give yet without learning about trigonometry first.

4.1 Basic trigonometric identities

Here I assume that the reader is already familiar with the basic trigonometric identities and their physical interpretation; therefore, I shall just list them out.

Function	Domain	Range
$\sin(\theta)$	$(-\infty,\infty)$	$[-1,1]$
$\cos(\theta)$	$(-\infty,\infty)$	$[-1,1]$
$\tan(\theta)$	$(-\infty,\infty)-\{\theta \vert \cos\theta=0\}$	$(-\infty,\infty)$
$\csc(\theta)$	$(-\infty,\infty)-\{\theta \vert \sin\theta=0\}$	$(-\infty,1]\cup[1,\infty)$
$\sec(\theta)$	$(-\infty,\infty)-\{\theta \vert \cos\theta=0\}$	$(-\infty,1]\cup[1,\infty)$
$\cot(\theta)$	$(-\infty,\infty)-\{\theta \vert \sin\theta=0\}$	$(-\infty,\infty)$
$\arcsin(\theta)$	$[-1,1]$	$[-\pi/2\,,\pi/2]$
$\arccos(\theta)$	$[-1,1]$	$[0,\pi]$
$\arctan(\theta)$	$(-\infty,\infty)$	$(-\pi/2\,,\pi/2)$
$\mathrm{arccot}(\theta)$	$(-\infty,\infty)$	$(0,\pi)$
$\mathrm{arcsec}(\theta)$	$(-\infty,1]\cup[-1,\infty)$	$[0,\pi/2)\cup(\pi/2\,,\pi]$
$\mathrm{arccsc}(\theta)$	$(-\infty,1]\cup[-1,\infty)$	$[-\pi/2\,,0)\cup(0,\pi/2]$

TABLE 4.1 | TABLE OF DOMAINS AND RANGE OF TRIGONOMETRIC FUNCTIONS

Pythagoras's identities:

$$\sin^2(\theta)+\cos^2(\theta)=1,\quad 1+\tan^2(\theta)=\sec^2(\theta),\quad 1+\cot^2(\theta)=\csc^2(\theta).$$

Angles addition:

$$\sin(A+B)=\sin A\cos B+\sin B\cos A,$$

$$\sin(A-B)=\sin A\cos B-\sin B\cos A,$$

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

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

Double angles formulas:

$$\sin(2\theta)=2\sin\theta\cos\theta,$$

$$\cos(2\theta) = \cos^2(\theta) - \sin^2(\theta) = 1 - 2\sin^2(\theta) = 2\cos^2(\theta) - 1,$$
$$\tan(2\theta) = \frac{2\tan\theta}{1 - \tan^2(\theta)}.$$

Triple angles formulas:

$$\sin(3\theta) = 3\sin(\theta) - 4\sin^3(\theta),$$
$$\cos(3\theta) = 4\cos^3(\theta) - 3\cos(\theta).$$

Half angles formulas:

$$\sin\left(\frac{\theta}{2}\right) = \sqrt{\frac{1-\cos\theta}{2}},$$
$$\cos\left(\frac{\theta}{2}\right) = \sqrt{\frac{1+\cos\theta}{2}},$$
$$\tan\left(\frac{\theta}{2}\right) = \sqrt{\frac{1-\cos\theta}{1+\cos\theta}}.$$

4.2 A visualization on sine and cosine

Sine and cosine are oscillating waves

4.3 Derivatives and integrals of basic trigonometric functions

I shall start with a simple example: what's the derivative of the function sine? You might turn yourself to the naive definition of derivative. But actually, if you look at the graph, you can actually figure out the derivatives yourself without needing the definition of derivative!

Illustrated in fig. 4.1, the slope at $x = 0$ can be approximated by a line $y = 1x + 0$: the derivative at 0 is 1. At $x = \pi/2$, the sine curve reaches its plateau: the derivative is 1. At $x = \pi$ it goes down by the slope -1, and

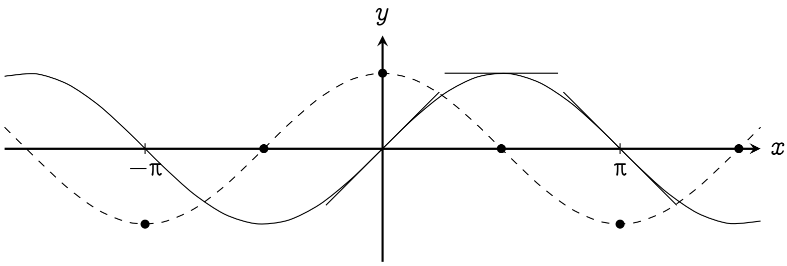


FIG. 4.1 | THE RELATION BETWEEN SINE AND COSINE

it *oscillates* like this so on and so forth. And it does this *smoothly*. The function that we know of today that does this oscillatory motion smoothly is clearly a cosine. So with the graph, you can already guess that the derivative of sine must be cosine.

We shall prove this by using **??**: the naive definition of derivative, and the angles addition formula.

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}\sin(x) &= \lim_{h\to 0}\frac{\sin(x+h)-\sin(x)}{h}\\ &= \lim_{h\to 0}\frac{\sin(x)\cos(h)+\sin(h)\cos(x)-\sin(x)}{h}\\ &= \lim_{h\to 0}\frac{\sin(x)(\cos(h)-1)+\sin(h)\cos(x)}{h}.\end{aligned}$$

When $h \to 0$, $\cos(h) \to 1$, thus $\cos(h)-1=0$. And, $\sin(h) \approx h$ for very small h. Therefore,

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

If I were to ask you again what the derivative of $\cos(x)$ is, well you can look at the graph again and instantly recognize that it's a $-\sin(x)$, and we could prove that by using the definition of derivative once again. But that would be a bit tiresome. Therefore, we can use the knowledge of cofunctions and the chain rule to help us.

$$\frac{\mathrm{d}}{\mathrm{d}x}\cos(x) = \frac{\mathrm{d}}{\mathrm{d}x}\sin(\pi/4 - x) \tag{4.1}$$

Then, let $u = \pi/4 - x$:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\cos(x) &= \frac{\mathrm{d}}{\mathrm{d}x}\sin(u) \\
&= \frac{\mathrm{d}}{\mathrm{d}u}\sin(u)\frac{\mathrm{d}u}{\mathrm{d}x} \\
&= \cos(u)\frac{\mathrm{d}}{\mathrm{d}x}(\pi/4 - x) \\
&= \cos(\pi/4 - x)(-1) = -\sin(x).
\end{aligned}$$

Voila! The derivative of cosine is the negative of sine.

If we take the derivative of $-\sin(x)$ again, it's $-\cos(x)$[1], and of $-\cos(x)$, $\sin(x)$. Thus, the derivatives of $\sin(x)$ forms a four part loop:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\sin(x) &= \cos(x), \\
\frac{\mathrm{d}}{\mathrm{d}x}\cos(x) &= -\sin(x), \\
\frac{\mathrm{d}}{\mathrm{d}x}(-\sin(x)) &= -\cos(x), \\
\frac{\mathrm{d}}{\mathrm{d}x}(-\cos(x)) &= \sin(x),
\end{aligned}$$

and the second derivative of sine is itself times a constant.

4.4 Finding the power series expansion of sine and cosine

By using the fact that the second derivative of sine is itself times a constant, one can also extract the power series expansion of sine just like what we did in section 3.6.1. We begin by letting

$$\sin(x) = a_0x^0 + a_1x^1 + a_2x^2 + a_3x^3 + \dots . \tag{4.2}$$

[1] Refer back to eq. (3.6)

Then,[2]

$$\frac{\mathrm{d}^2}{\mathrm{d}x^2}\sin(x) = a_2\frac{2!}{0!}x^0 + a_3\frac{3!}{1!}x^1 + a_4\frac{4!}{2!}x^2 + \dots . \tag{4.3}$$

The second derivative of sine must be the negative of sine, therefore

$$-a_0x^0 - a_1x^1 - a_2x^2 - a_3x^3 - \dots = a_2\frac{2!}{0!}x^0 + a_3\frac{3!}{1!}x^1 + a_4\frac{4!}{2!}x^2 + \dots$$

Matching the coefficient will give rise to a set of equations:

$$\begin{aligned} -a_0 &= \frac{2!}{0!}a_2, & -a_1 &= \frac{3!}{1!}a_3, \\ -a_2 &= \frac{4!}{2!}a_4, & -a_3 &= \frac{5!}{3!}a_5, \\ & \vdots \end{aligned}$$

Or generally,

$$-a_n = \frac{(n+2)!}{n!}a_{n+2}.$$

I've intentionally write the set of equations above in two separate columns because each column is independent of one another. Here, we have to find both a_0 and a_1 that satisfies the equation.

Just like in the exponential case, when plugged $x = 0$ in, the power series is left with just a_0, and $\sin(0) = 0$; therefore, $a_0 = 0$. The terms following a_0, namely the even numbered terms, also dissapears as a consequence, leaving us with the *odd* numbered terms:

$$\sin(x) = a_1x + a_3x^3 + a_5x^5 + \dots .$$

Now this might be a bit tricky, because there's virtually no way to get a_1 out by plugging anything in. Here, the small angle approximation will come in handy. Notice that we want our power series to *approach* sine of x with infinitely many terms. Truncating the series will obviously give

[2] $0! = 1$.

us a nice approximation to sine. Here is where the small angle approximation comes in handy. Let's truncate the series at the very first term.

$$\sin(x) \approx a_1 x.$$

This must be an approximation for sine. Which by the small angle approximation, $\sin(x) \approx x$ where x is small. Therefore, a_1 must be equal to one.

The rest of the equation is left for the reader as an exercise. Just take $a_1 = 1$ and go through the set of equations mentioned above. Then, you'll get that

$$\sin(x) = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \frac{1}{7!}x^7 + \ldots, \tag{4.4}$$

which shall justify my statement of eq. (3.17).

Now, to find the power series of cosine, you just have to differentiate eq. (4.4) and get

$$\cos(x) = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \frac{1}{6!}x^6 + \ldots.$$

As said, truncating these polynomials will actually get you the approximation for sine and cosine. I encourage you to take out any device that can plot graphs and put the first few terms in for both sine and cosine. In case that's not available, I illustrated the power series expansion of sine up to the fifth term for the reader in fig. 4.2.

4.5 Derivatives of other trigonometric functions

Note that these wouldn't be used that often until we reach chapter 7

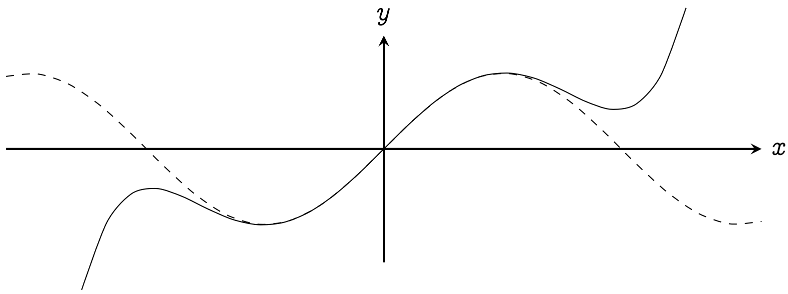


FIG. 4.2 | THE GRAPH OF SINE (DOTTED) PLOTTED WITH THE POWER SERIES EXPANSION OF SINE TO THE FIFTH TERM.

4.6 Derivatives and integrals of inverse trigonometric functions

4.7 Newton's law in polar coordinates

4.8 Newton's law in cylindrical coordinates

4.9 Harmonic addition theorem

In this section I shall extend on the trigonometric identities a bit wit the **harmonic addition theorem** which states that it's always possible to write a sum of sinusoids with the same angular speed, $a\sin(\omega\theta) + b\cos(\omega\theta)$, as a single sinusoid $c\cos(\theta+\phi)$. I'd love if there's a clean and nice geometrical interpretation for this, but apparently there isn't.[3] So, I shall just state the theorem with the proof below.

[3] I still have much questions with this proof here, e.g., how does this theorem goes with the phasors diagram?

Theorem 3: Harmonic addition theorem

The sum of sines and cosines with the same angular frequency ω can always be written as a sine function.

$$a\sin\left(\omega\theta\right) + b\cos\left(\omega\theta\right) = \operatorname{sgn}(b)\sqrt{a^2+b^2}\cos\left(\theta + \arctan\left(-\frac{a}{b}\right)\right).$$

where the sign function sgn is defined as

$$\operatorname{sgn}(x) = \begin{cases} -1 & \text{if} \quad x < 0 \\ 0 & \text{if} \quad x = 0 \\ 1 & \text{if} \quad x > 0 \end{cases}$$

Proof. First, notice that $c\cos\theta\cos\phi - c\sin\theta\sin\phi$. $\square$

CHAPTER 5

Significance of calculus

Abstract

I've massively restructured the contents of this chapter from the normal calculus textbook. Earlier in section 1.7, we see that calculus have much significance in kinematics. We'll be discussing about that later in chapter 9; however, I put a few of the worked out kinematics problems in the beginning of this chapter. After that, we'll be applying calculus in various other problems. Including

1. Higher dimensional quantities: area and volumes
2. Optimization
3. Root finding algorithm

Also, I will mention various techniques of integration needed along the way. Mainly, substitution of variables.

Prerequisites: *Basic derivatives and integrals (chapter 3), free body diagram writing*

5.1 Newton's fluxion notation

Before moving to further examples in kinematics, I'd like to introduce another notation called the **Newton's fluxion notation** or, the **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 $\dot{x}$.

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 Further examples of calculus in physics and kinematics

5.2.1 Block sliding down a ramp with friction

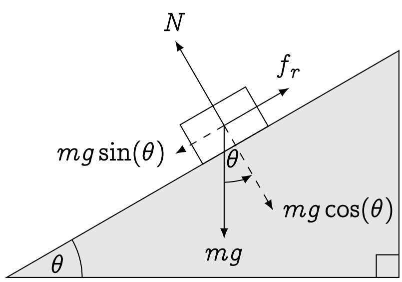


FIG. 5.1 | A BLOCK SLIDING DOWN A ROUGH RAMP (I.E. WITH FRICTION) ANGLED θ RELATIVE TO THE GROUND

Illustrated in fig. 5.1, 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 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

$$F_n = mg\sin(\theta) - f_r{}^{2}.$$

By the Newton's second law,

$$\begin{aligned} m\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right) &= mg\sin(\theta) - f_r \\ \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right) &= g\sin(\theta) - \frac{f_r}{m}. \end{aligned}$$

$g\sin(\theta) - f_r/m$ doesn't change with time, let's name it κ. The equation then becomes

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right) &= \kappa \qquad (5.1) \\ \int \mathrm{d}\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right) &= \int \kappa \, \mathrm{d}t \\ \frac{\mathrm{d}x}{\mathrm{d}t} &= \kappa t \\ \int \mathrm{d}x &= \int \kappa t \, \mathrm{d}t \\ x(t) &= \left(g\sin(\theta) - \frac{f_r}{m}\right)\frac{t^2}{2}. \qquad (5.2) \end{aligned}$$

[1] Just decompose mg into its parallel and perpendicular axis. Using basic trigonometry is enough.

[2] F_n for F_{net} or, total force.

which *is* our equation of motion. Notice, eq. (5.1) 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.

5.2.2 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 section 3.6.1):

$$f_r(\mathbf{v}) = a_0 + a_1\mathbf{v} + a_2\mathbf{v}^2 + a_3\mathbf{v}^3 + \dots .$$

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

$$m\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right) = a_1\frac{\mathrm{d}x}{\mathrm{d}t} + a_2\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right)^2 .$$

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

$$\frac{\mathrm{d}\dot{x}}{\mathrm{d}t} = p\dot{x} + q\dot{x}^2. \tag{5.3}$$

It is obvious that both p and q must be negative, otherwise the object would accelerate forward with the velocity. Frankly, eq. (5.3) is not possible to solve using the techniques that we have now. I'll revisit this exact differential equation later in section 8.1. 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 (5.3) 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.

$$\begin{aligned} \int \frac{1}{\dot{x}}\,\mathrm{d}\dot{x} &= p\int \mathrm{d}t \\ \ln(\dot{x}) + C &= pt. \end{aligned} \tag{5.4}$$

To find out what this integration constant should be, we have to use the initial condition $t = 0 \implies \dot{x} = v_0$.

$$\begin{aligned} \ln(v_0) + C &= 0 \\ C &= -\ln(v_0). \end{aligned}$$

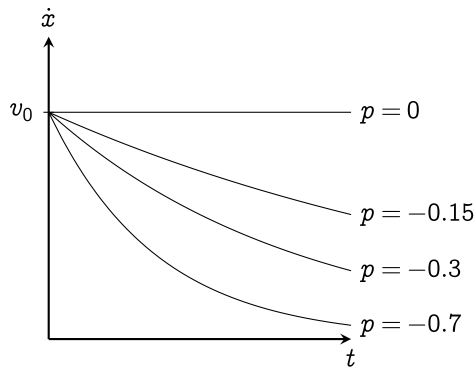


FIG. 5.2 | PLOT OF EQ. (5.5) WITH VARIOUS p.

Equation (5.4) then becomes

$$
\begin{aligned}
\ln(\dot{x}) - \ln(v_0) &= pt \\
\dot{x} &= v_0 \mathrm{e}^{pt}. \qquad (5.5)
\end{aligned}
$$

Figure 5.2 plots eq. (5.5) with various p. Notice that when $p = 0$, eq. (5.5) 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. (5.5) as

$$\frac{\mathrm{d}x}{\mathrm{d}t} = v_0 \mathrm{e}^{pt}.$$

Then,

$$
\begin{aligned}
\mathrm{d}x &= v_0 \mathrm{e}^{pt}\, \mathrm{d}t \qquad (5.6) \\
x + C_1 &= v_0 \int \mathrm{e}^{pt}\, \mathrm{d}t\,. \qquad (5.7)
\end{aligned}
$$

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 section 7.1. 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. (5.7), 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. (5.7), it reads

$$\begin{aligned} x + C_1 &= v_0 \int \mathrm{e}^{u} \left(\frac{1}{p}\,\mathrm{d}u \right) \\ x + C_1 &= \frac{v_0}{p}\mathrm{e}^{pt}. \end{aligned} \tag{5.8}$$

We also have to take care of the C_1. Using $t = 0 \implies x = x_0$,

$$\begin{aligned} x_0 + C_1 &= \frac{v_0}{p}\mathrm{e}^{p\cdot(0)} \\ C_1 &= \frac{v_0}{p} - x_0. \end{aligned}$$

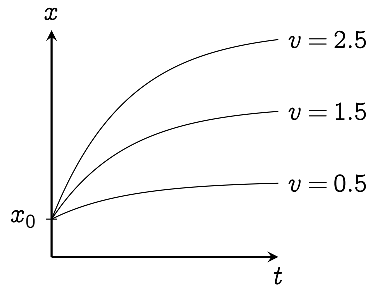
Then, plug it into eq. (5.8). The equation becomes

$$\begin{aligned} x + \frac{v_0}{p} - x_0 &= \frac{v_0}{p}(\mathrm{e}^{pt}) \\ x &= x_0 + \frac{v_0}{p}(\mathrm{e}^{pt} - 1). \end{aligned} \tag{5.9}$$

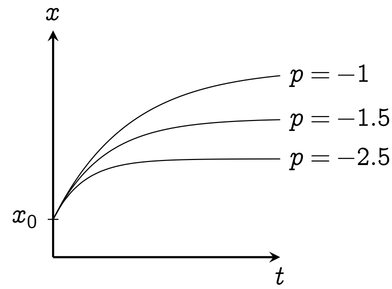
I've plotted eq. (5.9) with varying v in fig. 5.3a and varying p in fig. 5.3b. The graph should match with what you expect intuitively: higher v will get you further, and the less the drag, the further you'll get.

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. (5.9), we get

$$\lim_{t\to\infty} x(t) = \lim_{t\to\infty} \left(x_0 + \frac{v_0}{p}(e^{pt} - 1) \right).$$



(A) WITH VARYING v, SETTING $p=-1$



(B) WITH VARYING p, SETTING $v=2$

FIG. 5.3 | PLOT OF EQ. (5.9), SETTING $x_0 = 0.5$.

$$= x_0 + \frac{v_0}{p} \cancel{\lim_{t\to 0}\left(e^{pt}\right)}^{0} - \frac{v_0}{p} = x_0 - \frac{v_0}{p},$$

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 (5.3) simplifies to

$$\frac{\mathrm{d}\dot{x}}{\mathrm{d}t} = q\dot{x}^2$$
$$\int \frac{1}{\dot{x}^2}\,\mathrm{d}\dot{x} = q\int \mathrm{d}t \text{ [3]}$$
$$-\frac{1}{\dot{x}} + C = qt.$$

Taking care of C: use $t=0 \implies \dot{x} = v_0$.

$$-\frac{1}{v_0} + C = q\cdot 0$$
$$C = \frac{1}{v_0}.$$

Therefore,

$$-\frac{1}{\dot{x}} + \frac{1}{v_0} = qt$$
$$\frac{\mathrm{d}t}{\mathrm{d}x} = \frac{1-qv_0t}{v_0}$$
$$x = v_0\int \frac{1}{1-qv_0t}\,\mathrm{d}t\,. \quad (5.10)$$

The structure of the integral is similar to $1/t$. Therefore, let $u = 1 - qv_0t$. Then,

$$\frac{\mathrm{d}u}{\mathrm{d}u} = \frac{\mathrm{d}}{\mathrm{d}u}(1-qv_0t)$$
$$1 = -qv_0\frac{\mathrm{d}t}{\mathrm{d}u}$$
$$\mathrm{d}t = -\frac{1}{qv_0}\,\mathrm{d}u\,.$$

Substitute $u = 1 - qv_0 t$ and $\mathrm{d}t = -\frac{1}{qv_0}\,\mathrm{d}u$ into eq. (5.10):

$$\begin{aligned} x &= v_0 \int \frac{1}{u}\left(-\frac{1}{qv_0}\,\mathrm{d}u\right) \\ &= -\frac{1}{q}\int \frac{1}{u}\,\mathrm{d}u \\ x(t) &= -\frac{1}{q}\ln\left(1 - qv_0 t\right) + C_1. \end{aligned} \tag{5.11}$$

Taking care of C_1: use $t \to 0 \implies x(t) = x_0$.

$$\begin{aligned} x_0 &= -\frac{1}{q}\ln\left(1 - qv_0 \cdot (0)\right) + C_1 \\ x_0 &= C_1 \end{aligned}$$

Plug this into eq. (5.11) to get the final answer:

$$x(t) = -\frac{1}{q}\ln\left(1 - qv_0 t\right) + x_0.$$

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.

5.2.3 Time of meteor collision from great height

Illustrated in fig. 5.4[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

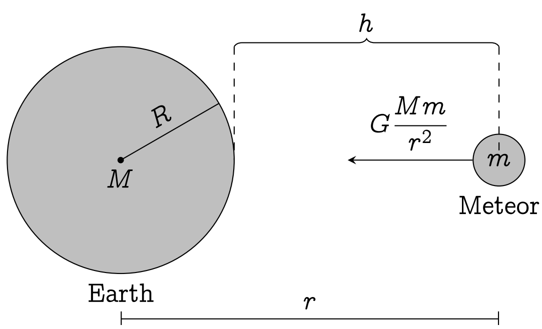


FIG. 5.4 | AN ILLUSTRATION OF A METEOR FALLING TO THE EARTH FROM HEIGHT h.

mass m, falling from height h above the ground. The Earth has mass M,

[4]Not to scale.

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 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 $\kappa = GM$

$$\frac{\kappa}{r^2} = \frac{\mathrm{d}\dot{r}}{\mathrm{d}t}. \tag{5.12}$$

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 these variables. Here, I shall introduce a technique for solving this kind of differential equation. With the chain rule,

$$\frac{\mathrm{d}\dot{r}}{\mathrm{d}t} = \frac{\mathrm{d}\dot{r}}{\mathrm{d}r} \times \frac{\mathrm{d}r}{\mathrm{d}t} = \dot{r}\frac{\mathrm{d}\dot{r}}{\mathrm{d}r}:$$

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 $\mathrm{d}\dot{r}/\mathrm{d}r$ as the velocity at any given distance away

[5] Notice that here, r also changes with time.

from the Earth. Plug this into eq. (5.12), we get

$$\begin{aligned}\frac{\kappa}{r^2} &= \dot{r}\frac{\mathrm{d}\dot{r}}{\mathrm{d}r}\\ \kappa\int\frac{1}{r^2}\,\mathrm{d}r &= \int\dot{r}\,\mathrm{d}\dot{r}\\ \frac{\kappa}{r}+C &= \frac{\dot{r}^2}{2} \qquad (5.13)\end{aligned}$$

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. (5.13), 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.

5.2.4 One-dimensional simple harmonic motion

A harmonic oscillator is described by an object that's under a force $F(x) = -kx$, which is a function of position. Here, we can use Newton's second law straightaway:

$$m\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right) = -kx \tag{5.14}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathrm{d}x}{\mathrm{d}t}\right) = -\frac{k}{m}x. \tag{5.15}$$

There are actually three ways of approaching this, which I shall get you through all three.

Ansatz method of solving differential equations

The first one is called the *ansatz* method, which is commonly taught in the MIT university. Ansatz means to make assumptions; this method assumes the solution of the differential equation then finalize it later. In this case, you have to ask yourself what function when differentiated twice w.r.t. time gives the function itself times a constant? Well, there are two functions which satisfies this: $\sin(t)$ and $\cos(t)$.

Notice that this differential equation is *linear*. That means, if $f(t)$ is a solution, and $g(t)$ is a solution, then $f(x)+g(x)$ is also a solution. Therefore, our ansatz might look something like $\sin(t)+\cos(t)$. Our ansatz when differentiated twice must have a constant times itself. A pretty general ansatz that one might think of is $A\sin(\omega t) + B\cos(\omega t)$.

There's a reason why I used ω as a variable. ω suggests the angular speed of the $\sin(t)$ and $\cos(t)$. It will be evident later why it's the angular speed.

Now, with the ansatz in place, it's time to find A, B, and ω. We can do that by just substituting the ansatz into eq. (5.15):

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathrm{d}}{\mathrm{d}t}(A\sin(\omega t) + B\cos(\omega t))\right) = -\frac{k}{m}\left(A\sin(\omega t) + B\cos(\omega t)\right).$$

We can then use the chain rule to the L.H.S. It's left as an exercise to the reader to verify that this is true.

$$
\begin{aligned}
-A\omega^2 \sin(\omega t) - B\omega^2 \cos(\omega t) &= -\frac{k}{m} A \sin(\omega t) - \frac{k}{m} B \sin(\omega t) \\
A\omega^2 \sin(\omega t) + B\omega^2 \cos(\omega t) &= \frac{k}{m} A \sin(\omega t) + \frac{k}{m} B \sin(\omega t).
\end{aligned}
$$

Here, we can match the coefficient and separate our equation into two parts:

$$
\begin{aligned}
A\omega^2 \sin(\omega t) &= \frac{k}{m} A \sin(\omega t), \\
B\omega^2 \cos(\omega t) &= \frac{k}{m} B \cos(\omega t).
\end{aligned}
$$

Both equations say that $(k/m)^{1/2}$ must be equal to ω. However, notice that both equations does not say anything about A, or B. That is simply because we can choose A and B ourselves; it's a free parameter. Our final solution is then

$$
x(t) = A \sin\left(\sqrt{\frac{k}{m}} t\right) + B \cos\left(\sqrt{\frac{k}{m}} t\right).
$$

The series expansion method

5.2.5 Harmonic motion in two-dimensions: the pendulum

The repeated integration method

5.2.6 Damped harmonic motion

5.3 Conservation laws

5.4 Volumes of solids

5.5 The amount of real zeroes in a cubic equation

5.6 Optimization problems

5.7 Principle of least action

5.8 Tangent to a curve

5.9 Newton's root finding algorithm

CHAPTER 6

Series

6.1 Sequences

6.2 Geometric series: zeno's paradox

6.3 Convergence and divergence

6.3.1 Absolute convergence

6.4 Convergence test

6.5 Harmonic motion: the need of functions approximation

6.6 Taylor series

CHAPTER 7

Basic techniques of integration

Abstract

This chapter explores various techniques you can use to integrate stuffs. You still have to keep in mind that even though this chapter is very abstracted away, *integrals still have its geometric meaning: area under the graph.* This geometric interpretation is going to somewhat appear quite often throughout the chapter, so be aware of it.

Also, some integrals are definite: they have clear boundaries of integrations. You also have to keep track of what variable is bounded by that bound, otherwise you might end up evaluating incorrectly.

Note from the writer: As a calculus teacher myself, I usually don't teach this topic but leave it as an exercise. Techniques of integrations come naturally with much experiences. Therefore, I list a table of integrations at the end of the chapter for the physicists out there.

7.1 Change of variables

Change of variables, or integration by substitution is a method to convert an unknown integrals into a more familiar one. Consider $\int x\mathrm{e}^(x^2)\,\mathrm{d}x$. This integral might seem daunting at first but, notice that the derivative of x^2 is exactly x. I shall introduce a new variable u where $u = x^2$. If we take the derivative w.r.t. x on both sides,

$$\frac{\mathrm{d}u}{\mathrm{d}x} = 2x$$
$$\frac{\mathrm{d}}{\mathrm{d}u} = 2x\,\mathrm{d}x\,.$$

7.2 Integrals of inverse functions

7.3 Integrals of symmetric functions

7.3.1 Even and odd functions

7.3.2 Functions with axial and spherical symmetry

7.4 Integration by part

7.5 Recurrence relations and reduction formulas

7.6 Working with complex numbers and exponentials

7.7 Trigonometric substitution

7.8 Rational functions

7.8.1 Partial fraction decomposition

7.8.2 Quadratic under radicals

7.8.3 Ostrogradski method

7.9 Feynman's trick for integration

7.10 Polar integration

7.11 Gaussian integral

7.12 Cauchy's formula for repeated integrations

7.13 Numerical integration

CHAPTER 8

Advanced techniques of integration

Abstract

This is the chapter that explores more advanced techniques of integrations. You could go on with your life skipping this chapter and it'd still be fine. However, for the curious, there are some very interesting maths in here, so be sure to check this chapter out before going on to the next part of the book.

8.1 Rational functions

8.1.1 Quadratics denominator

An integral in the form

$$\int \frac{1}{ax^2 + bx + c} \, \mathrm{d}x \tag{8.1}$$

can be evaluated by completing the squares:

$$ax^2 + bx + c = a\left(x + \frac{b}{2a}\right)^2 + \left(c - \frac{b^2}{4a}\right) \tag{8.2}$$

8.1.2 Denominator quadratics under radicals

8.1.3 Feynman's trick for integration

Formerly known as Leibniz's rule but popularized by Richard Feynman, is a very powerful integral technique that allows you to differentiate the function in the integral sign. Sadly, it only works for definite integrals. However, it's still a very nice trick to have just in case.

The idea is: integrals and derivatives can be swapped in

Part II

The applications

CHAPTER 9

Ordinary differential equations I: basic concepts

9.1 Basic example: projectile motion

Part III

The extensions

Part IV

The fundamentals, reimagined: real analysis

CHAPTER 10

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.

10.1 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.

10.2 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.

$$\forall X \forall Y [\forall z (z \in X \iff z \in X) \implies X = Y]. \tag{10.1}$$

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*.

$$\forall x[x \neq \emptyset \implies \exists y((y \in x) \wedge (y \cup x) = \emptyset)] \tag{10.2}$$

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]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).

[2]Wikipedia contributors, *Summation: powers and logarithm of arithmetic progressions*, (Feb. 4, 2024) https://en.wikipedia.org/wiki/Summation#Powers_and_logarithm_of_arithmetic_progressions.