Mathematical Lecture 2

Anson Poon 5G 23

October 2025

Abstract

These series of documents aim to equip you with the tools needed for DSE M1/2, Physics, or advanced math and physics for those who are interested.

Please note that

1. The formatting and notation of this document might not be standard in public examinations.
2. Some tools are not allowed in internal and/or public exams.
3. Please don't kill me if I made any typos in this document. I will fix them upon request.
4. I would like to introduce the tools instead of drilling in the rigor. Rigorous proofs or definitions can be found at the end of the document. An asterisk (*) indicates that the rigorous construction is given at the end of the document.
5. The parts that are out-syllabus are marked with !!!.

1 Introduction

This document will focus on differentiation. Again, for the sake of explanation and visualization, (unfortunately) rigor has to be set aside for now.

Before I start, I would like to say that $\frac{dy}{dx}$ is not a fraction in the most rigorous sense. However, as the Leibniz notation makes theorems like chain rule easy to understand, you can treat them as fractions for now.

In the high school physics class, we are expected to solve straight line relations between displacement and time, or velocity and time, but mentions nothing about non-constant relations. This document will hopefully equip you with the idea of differentiation, in which a lot of modern math and physics is built upon.

2 Slope of a straight line

I would like to recap the idea of the slope of a straight line as this is important for the understanding of the concept of differential calculus.

Given two points $P(x_1, y_1), Q(x_2, y_2)$ in the 2D plane, the *slope* of the line PQ is defined as:

$$m_{PQ} := \frac{y_2 - y_1}{x_2 - x_1}$$

The slope gives an intuitive sense of the notion of "rate of growth": for a line with slope 2, an increment of x will result in double the increment of y, as you should learnt in F2 math (or F3? idk). This has a rather clear interpretation of "y grows twice as fast as x at every point on the line".

A few more observations could be done before heading to the next section.

1. While this definition works for straight lines, this doesn't work for curved graphs. Take for example $y = x^2$ and a point $P(2, 4)$ on the graph. Choose another point $Q(t, t^2)$ and perform the slope calculation on PQ, the slope would be

$$m_{PQ} = \frac{t^2 - 4}{t - 2} = t + 2 \qquad\qquad --- (2.1)$$

   Note that the denominator shows that $t \neq 2$ as Q collapses into P, i.e. when $t = 2$, $Q = P$, which no meaningful notion of slope could be discussed.

2. Although m_{PQ} depends on Q, we are still free to move Q around, e.g. setting $t = 2.5, 10, 10^6$ and see how the line segment PQ matches with the graph, and more importantly, see how the slope (rate of change) of the line segment PQ matches with the graph's rate of change at P.

3 Differentiation

Differentiation was invented to generalize the notion of slopes to curve graphs, in the sense that the *slope* of the graph $y = 3x + 3$ is constant and well-defined for all inputs $x \in \mathbb{R}$, but not the slope of the graph $y = x^2$ at $x = 3$, for example. It could be seen that the graph is steeper as x increases. It is flatter near $x = 0$, and is steeper as x increases.

A reasonable attempt might be to choose points near $(3, 9)$ on the graph, for example $(4, 16)$ and $(2, 4)$. However, you could see that if you join $(3, 9)$ with either point, the slope doesn't quite match the slope of the graph at $x = 3$.

The process of *differentiation* has different interpretations due to mathematicians caring too much about functions that work weirdly and physicists who doesn't care at all. Again, I will use the physicist's definition for now.

3.1 Definition

The *derivative* of a function $f : U \to V$, where U, V are closed intervals in $\mathbb{R}$, is defined as such:

$$y' = \frac{dy}{dx} = \frac{d}{dx}(y) := \frac{\text{small change in } y}{\text{small change in } x}$$

By small, I mean an *infinitesimally* small change. Written out in symbols, (also denote $y := f(x)$): between

$$\frac{d}{dx}f(x) = \frac{dy}{dx} = \frac{f(x + \Delta x) - f(x)}{\Delta x} \qquad\qquad --- (3.1.1)$$

where Δx is very small (hence the small changes in x and y values respectively).

In the context of the previous example with two points P, Q on the graph $y = x^2$, the derivative of the function at $x = 2$ can be obtained through slightly modifying the slope expression in (2.1).
By letting Q to be defined as $2 + \Delta x$ instead (the point of focus is $x = 2$, so it makes sense to define Q based on the distance from P), the equation becomes:

$$m_{PQ} = \frac{(2+\Delta x)^2 - 4}{\Delta x} \qquad --- (3.1.2)$$

However, this is not entirely the definition of the derivative. The derivative is defined as the quotient above when the difference between x-coordinate of P, Q are close. Hence, we need to notate that we are taking Δx to be close to 0. This process is called taking a limit. The complete definition is as follows:

$$\frac{dy}{dx}\Big|_{x=2} = \lim_{\Delta x \to 0} \frac{(2+\Delta x)^2 - 2^2}{\Delta x} \qquad --- (3.1.3)$$

where $\lim_{x \to 0}$ indicates that the quotient is taken as the *limit** when Δx is infinitely close but not zero, and $|_{x=2}$ indicates that we are calculating the slope of the curve at $x = 2$. This "small but not zero" idea is useful for further understanding of differential calculus.

However, the expression in (3.1.3) is not entirely satisfying. By looking at the numerator and the denominator, you can see that they "tend to" (i.e. gets infinitely close to) 0. In some sense, you cannot extract any useful information from the RHS of (3.1.3).
Observe that you can expand out the numerator on the RHS as such:

$$\frac{dy}{dx}\Big|_{x=2} = \lim_{\Delta x \to 0} \frac{4\Delta x + \Delta x^2}{\Delta x} = \lim_{\Delta x \to 0} 4 + \Delta x$$

The expression $\lim_{\Delta x \to 0} 4 + \Delta x$ means "$4 + \Delta x$ evaluated when Δx is close to 0". Hence, the RHS evaluates to $4 + 0 = 4$.

Note that this doesn't work just for $x = 2$. For the graph $y = x^2$, the slope at any x value is given by

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{(x+\Delta x)^2 - x^2}{\Delta x} = \lim_{\Delta x \to 0} 2x + \Delta x = 2x$$

in which the "2"s in the previous example are replaced by x. This shows that the slope of $y = x^2$ at x is $2x$.

3.1.1 Examples

I will gives some examples of common derivatives here. The proofs can be easily found online but I might include them in the future. Here, $\frac{d}{dx}f(x)$ denotes the derivative of the function f.

1. $\frac{d}{dx}x^n = nx^{n-1}$
2. $\frac{d}{dx}\sin x = \cos x$

3. $\frac{d}{dx} \cos x = -\sin x$

4. $\frac{d}{dx} b^x = b^x \ln b$, where $b > 0$ and ln denotes the natural logarithm.

5. $\frac{d}{dx} \ln x = \frac{1}{x}$

6. $\frac{d}{dx} \tan x = \sec^2 x$

You are highly recommended to search and go through their proofs, whether through online sources or your M2 textbook.

3.2 Rules

There are some so-called rules for differentiation. These include the addition, product, reciprocal and chain rule.

3.2.1 Addition Rule

For addition, the derivative of sums of functions are the sums of derivative of functions. In symbols, this writes:

$$\frac{d}{dx}(f(x) + g(x)) = f'(x) + g'(x) \qquad --- (3.2.1)$$

This should make intuitive sense. For example, if f, g have a rate of change of 3s^{-1} and 4s^{-1} respectively, they should have a combined rate of change of 7s^{-1}. This is an easy rule to remember, the derivative can be distributed through addition like multiplication ($3(1 + 2) = 3 \cdot 1 + 3 \cdot 2$, similar to $(f + g)' = f' + g'$). This is the most easy one to understand among the four rules. The following ones are more painful to make intuitive sense of.

Infinitesimal notation The following sections will use infinitesimal notation. Note that from our previous definitions,

$$y' = \frac{dy}{dx}$$

From this, we "slightly" abuse notation to obtain the following:

$$dy = y' dx \qquad --- (3.2.2)$$

These equations are both true but not in the way that you would think. However, I wish not to talk about the technical details here and will proceed to cry in the lack of mathematical rigor (for the sake of clarity). Interested ones who want to know why $\frac{dy}{dx}$ is not a fraction are encouraged to search it up (short answer: treating it as a fraction assumes that dy, dx work similarly with numbers, however $\frac{dy}{dx}$ is a limit as shown in the definition, which you can't always move numerators or denominators around).

To give credit where its due, this makes some sense. We are interested in the infinitesimal change of y (i.e. dy), when there is an infinitesimal change in x (i.e. dx). Hence, it would be sensible to express dy in terms of dx.

3.2.2 Product Rule

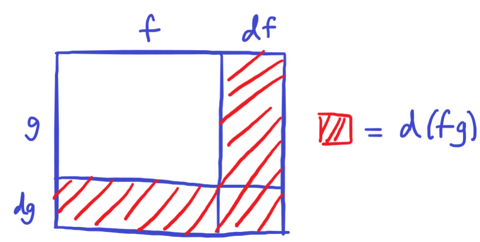
The product rule states that for two differentiable functions f, g, the derivative of their product (exists and) are given by

$$(fg)' = f'g + g'f \qquad --- (3.2.3)$$

or restated as

$$\frac{d}{dx}(f(x)g(x)) = f'(x)g(x) + g'(x)f(x) \qquad --- (3.2.4)$$

To understand this seemingly unintuitive rule, I would use the following diagram to illustrate geometrically.



As you can see, under a small change dx in the x value, f, g would also have a small change of df, dg respectively. Hence, the change in fg would be:

$$d(fg) = (f + df)(g + dg) - fg = fdg + gdf + dfdg \qquad --- (3.2.5)$$

where $d(fg)$ denotes a small change in fg. Notice that $dfdg$ is very small compared to fdg and gdf. We'll neglect the effects of this term, in the sense that $dfdg$ decreases much quicker than df and dg*. (I understand that this is not intuitive but please follow through for now. The justifications have been put at the end of the document.) Hence,

$$d(fg) = fdg + gdf \qquad --- (3.2.6)$$

Using (3.2.2), i.e. $d(\text{function}) = (\text{derivative})dx$, $d(fg) = (fg)'dx, df = f'dx, dg = g'dx$:

$$(fg)'dx = (fg')dx + (gf')dx = (f'g + g'f)dx$$

Hence, finally, the product rule can be obtained.
More than one infinitesimal (e.g. df, dx, dg) multiplied together will be taken as 0*.

3.2.3 Reciprocal rule

The derivative of reciprocal of functions are not as intuitive to find as it seems. A reasonable attempt would be

$$0 = d(1) = d(f \cdot \frac{1}{f}) = \frac{1}{f}df + fd(\frac{1}{f}) \qquad --- (3.2.7)$$

This is a step in the right direction (apparently). Expanding $df = f'dx$ and rearranging gives:

$$d(\frac{1}{f}) = \frac{-f'}{f^2}dx$$

Hence, using (3.2.2) on the LHS,

$$(\frac{1}{f})' = \frac{-f'}{f^2} \qquad ---(3.2.8)$$

This is called the *reciprocal rule.*

Quotient rule I despise the quotient rule for personal reasons. However, quotient rule propaganda has been spread. Those who wish to know can derive it as such:

$$d(\frac{f}{g}) = d(f \cdot \frac{1}{g}) = fd(\frac{1}{g}) + \frac{df}{g} \qquad ---(3.2.9)$$

Using the reciprocal rule,

$$d(\frac{f}{g}) = (\frac{-g'f}{g^2} + \frac{f'}{g})dx = \frac{f'g - g'f}{g^2}dx$$

Hence the quotient rule can be obtained:

$$(\frac{f}{g})' = \frac{f'g - g'f}{g^2} \qquad --(3.2.10)$$

3.2.4 Chain rule

The chain rule is used to evaluate the derivatives of composite functions such as $g(x) = f(u(x))$. This uses a trick.

By (3.2.2), $dg = f'(u)du$. This holds true even if u is a function. However, we wish to obtain a derivative with respect to x, not u.
Observe that $du = u'(x)dx$ by, again, (3.2.2). Combining gives:

$$dg = f'(u)du = f'(u(x))u'(x)dx$$

At the end, we can obtain the following:

$$g'dx = f'(u(x))u'(x)dx$$
$$g' = f'(u(x))u'(x)$$

As this rule, I think, is one of the most important rules of differential calculus, I'll give some examples below to demonstrate.

Examples

1. $d(\cos x^2)$, let $u = x^2$, $d(\cos x^2) = d(\cos u) = -\sin udu = -\sin u(2xdx) = -\sin x^2(2x)dx$
   Hence, $\frac{d}{dx}\cos x^2 = -2x\sin x^2$
2. $d(\sin\sqrt{2\ln x})$, let $u = 2\ln x$, $t = u^{0.5}$, $d(\sin\sqrt{2\ln x}) = d(\sin t) = \cos tdt = \cos t(0.5u^{-0.5})du = \cos t(0.5u^{-0.5})(2x^{-1})dx$. Substitute the u and t back in terms of x.
3. $d(e^{\cos x \cdot x^2})$, let $u = (\cos x \cdot x^2)$, then $d(e^{\cos x \cdot x^2}) = d(e^u) = e^u du = e^u((2x\cos x - x^2\sin x)dx)$
   Hence, $\frac{d}{dx}(e^{x^2\cos x}) = e^{x^2\cos x}(2x\cos x - x^2\sin x)$

This is getting repetitive and painful to type.

3.3 Some applications

In physics, differentiation is a very useful tool to describe rate of change, as we are often investigating changes of some parameters such as position or velocity with respect to time. Here are some examples of differentiation in physics:

1. Velocity $v = \frac{dx}{dt}$, where x is the position
2. Acceleration $a = \frac{dv}{dt}$
3. Angular velocity $\omega = \frac{d\theta}{dt}$, where θ is the angular displacement
4. Angular acceleration $\alpha = \frac{d\omega}{dt}$
5. Force $F = \frac{dp}{dt}$, where p is the momentum

4 Approximations of functions

4.1 First degree approximation

This section is useful for those who have a habit of not bringing their calculator to their math and/or physics quizzes and have to calculate square roots by hand.

Given a function $f(x)$ which we take to be differentiable, the slope of the tangent line at $(a, f(a))$ is $f'(a)$. Hence we can construct the point-slope form of the line:

$$L_a : y - f(a) = f'(a)(x - a)$$

This line approximates the values of f sufficiently close to a, hence a linear approximation. Take $\sqrt{25.4}$ for example, the true value is 5.03984126734. We can take the tangent line at $(x, y) = (25, 5)$ to get:

$$y = 5 + 0.1(x - 25)$$

Substituting $x = 25.4$,

$$y = 5 + 0.1(0.4) = 5.04$$

which squares to 25.4016, being very close to the true value of 5.0398. This is, in fact, the best line that gives an approximation of $\sqrt{x}$ around $x = 25$.

Some of you might notice the title and see the words "first-degree". This is because we get an approximation $f(x_1)$ through the line $y = f(x_0) + f'(x_0)(x_1 - x_0)$ while only considering a first derivative. However, the new value is a better approximation than the initial x_0 and a new line could be constructed to give a even better approximation.

4.2 Higher order approximation and Taylor series

Notice how we got the tangent line and found an approximation. There are two ways to improve the approximation. One is constructing more tangent lines and the other is constructing a series of the function.

4.2.1 Tangent lines and the Newton-Raphson Method

This method has two branches of approaches. The first is by taking the initial guess x_0 and getting x_1 by the linear approximation, and then using x_1 to get x_2 through another tangent line, and so on. This method works for sufficiently nice functions and initial values. There is a nicer way to do this.

Take a function $y = f(x)$. The approach now is to take an initial guess x_0, and constructing the tangent line $y - f(x_0) = f'(x_0)(x - x_0)$, but this time we're interested in the x-intercepts, as we are using the x-intercepts of the lines to approximate the root (i.e. x-intercept of the curve). We'll see how the x-intercepts are useful later

We'll get that

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

However, we could do this again to generate $L_1 : y - f(x_1) = f'(x_1)(x - x_1)$ to obtain

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}$$

where each x_n is a better approximation to x_{n-1}. In general, if you want to numerically approximate the roots of a function, you would set up a recursion

$$x_n = x_{n-1} - \frac{f(x_{n-1})}{f'(x_{n-1})}$$

and you'd find the roots of x if you chose x_0 carefully. The solution desired is $\lim_{n\to\infty} x_n$ (or x_ω if I'm lazy).

4.2.2 Taylor series !!!

It is a mathematical fact that "well-behaved" functions can be approximated by a series:

$$f(x) = a_0 + a_1 x + a_2 x^2 + ... = \sum_{k\in\mathbb{N}} a_k x^k$$

with infinitely many terms. This is included in the definition of the exponential function:

$$e^x := \sum_{k\in\mathbb{N}} \frac{x^k}{k!}$$

There is a way to find the coefficients of x^k. The amazing observation is that $x^k = 0$ when $x = 0$. We can exploit this fact by setting $x = 0$ on both sides to obtain:

$$f(0) = a_0$$

Now we've obtained the constant term we need. Then, we take the derivative on both sides to get

$$f'(x) = a_1 + 2a_2 x + 3a_2 x^2 + ...$$

Again, we can set $x = 0$ to get
$$f'(0) = a_1$$
In a similar fashion we could obtain the power series of a (analytic) function:
$$f(x) = \frac{f(0)}{0!} + \frac{f'(0)}{1!}x + \frac{f''(0)}{2!}x^2 + ... = \sum_{n\in\mathbb{N}} \frac{d^n}{dx^n} f(x)\bigg|_{x=0} \frac{x^n}{n!}$$
This is known as a *Maclaurin series*. The more general form:
$$f(x) = \sum_{n\in\mathbb{N}} \frac{d^n}{dx^n} f(x)\bigg|_{x=a} \frac{(x-a)^n}{n!}$$
is known as a Taylor series centered at a.

Some examples I'll give some examples of Maclaurin series here:
$$e^x = \sum_{n\in\mathbb{N}} \frac{d^n}{dx^n} e^x\bigg|_{x=0} \frac{x^n}{n!} = \sum_{n\in\mathbb{N}} e^x\bigg|_{x=0} \frac{x^n}{n!} = \sum_{n\in\mathbb{N}} \frac{x^n}{n!}$$
$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} ... = \sum_{n\in\mathbb{N}} \frac{d^n}{dx^n} \cos x\bigg|_{x=0} \frac{x^n}{n!} = \sum_{n\in\mathbb{N}} \frac{x^{2n}}{(2n)!}$$
$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} ... = \sum_{n\in\mathbb{N}} \frac{d^n}{dx^n} \sin x\bigg|_{x=0} \frac{x^n}{n!} = \sum_{n\in\mathbb{N}} \frac{x^{2n+1}}{(2n+1)!}$$
$$-\ln(1-x) = \sum_{n\in\mathbb{N}\setminus\{0\}} \frac{x^n}{n}$$
To approximate the function, compute the first few terms of the series. *Remark:* The x in the series of for sin and cos is in radians.

5 Appendix !!!

5.1 Limits

Limits describe how the function behaves near a point, whether it's defined at that point or not.
To illustrate this, consider the function:
$$y = \frac{x^2 - 16}{x - 4}$$
This function is defined almost everywhere and has a discontinuity at $x = 4$. (Input this graph in a graphing calculator and try to reason why there's a hole at $x = 4$)
However, you might look at the function and argue that it's very similar to $y = x + 4$. Indeed, factorizing the numerator yields:
$$y = \frac{(x-4)(x+4)}{x-4} = x + 4 \quad (x \neq 4)$$

You could reason that although the graph has a hole at $x = 4$, the value of y near $x = 4$ gets arbitrarily close to 8.

This beings us to the definition of the limit. To illustrate this in English instead of symbols, I'll introduce some open-ended questions.

Consider two functions:

$$f(x) = x^2,\ g(x) = \frac{x^2 - 4}{x - 2}$$

Here are the questions:

1. Draw two horizontal lines: $y = 4 + 1$ and $y = 4 - 1$. At the points where they intersect the graph, draw a vertical line to the x-axis. There is a box you can form from these two sets of lines. Can you deduce that the area of the box to be finite or not?
2. Draw the box again but with $y = 4 \pm 0.5$ instead. Does the area of the box increase or decrease?
3. Draw the box again but with $y = 4 \pm 0.25$ instead. Through the three boxes you've drawn, is the box converging to a point? If yes, then what is the point that it's converging to?

The point that the function converges to is the *limit* of the function at $x = 2$. Even if the function is not defined at $x = 2$, we can see that the box always contains $(2, 4)$. This could be intuitively interpreted as the function converging to a certain value as x approach 2. Hence, the limit of both functions are 4.

Formal definition I have actually secretly told you a simplified version of the formal definition process above (xd). In symbols,

$$\lim_{x \to a} f(x) = L \iff (\forall \epsilon > 0)(\exists \delta > 0)(\forall x \in \mathbb{R})(0 < |x - a| < \delta \implies |f(x) - L| < \epsilon)$$

This means, for all ranges of $f(x)$ I choose around L, if I can also find a finite range for x around a such that all x in the range is in the $f(x)$ range chosen. Or in human terms, the function is confined in the box for all ranges of f chosen.

5.2 $dx^n = 0$

In the previous section, I have stated that $dx^n = 0$ for $n \geq 1$. This could be also seen in $df dg = (f'dx)(g'dx) = (f'g')dx^2 = 0$
This could be seen in the derivative of $y = x^2$ as an example. Previously, I've shown that the derivative of x^2 is $2x$ through the following limit:

$$\frac{d}{dx}x^2 = \lim_{\Delta x \to 0} \frac{(x + \Delta x)^2 - x^2}{\Delta x} = \lim_{\Delta x \to 0} 2x + \Delta x$$

In the numerator, terms that have no Δx term are canceled, and the term with the power of Δx to be only exactly 1 is concerned, as higher order Δx's are taken to be 0 in the limit.
Also, by the notion that $\lim_{\Delta x \to 0} \Delta x = dx$, only the terms with dx^1 are concerned, while $dx^n = \lim \Delta x^n = 0$ for all $n \geq 2$.

5.3 Fraction?

Notice that previously when I introduced the derivative, two notations are used. This gives rise to the equality:

$$dy = \frac{dy}{dx}dx$$

So, is $\frac{dy}{dx}$ a fraction?
Short answer, no. The derivative is the limit of a ratio, not a ratio of limits.
Long answer, observe that the definition of the derivative is the limit:

$$\lim_{h\to 0} \frac{f(x+h) - f(x)}{h}$$

while dy and dx are defined as

$$dx = \lim_{h\to 0}((x+h) - x), dy = \lim_{h\to 0}(f(x+h) - f(x))$$

where $y = f(x)$. So, what $\frac{dy}{dx}$ actually means is:

$$\frac{dy}{dx} = \frac{\lim_{h\to 0}(f(x+h) - f(x))}{\lim_{t\to 0} t}$$

which is not necessarily equal to the actual definition, because limits are sometimes weird.