



مجموعة جامعة
الجميع الذكية
للتعليم القابضة

MIDOCEAN
UNIVERSITY
BY EVERYONE'S SMART UNIVERSITY

Mathematical Foundations For Data Science and AI

Lecture 01

AI102

Calculus; Introduction to Derivative

By
Dr. Khaled ELMenshawy

www.midocean.ae

Lecture

01

LEARNING OUTCOMES

- Understanding Number Theory and Variables.
- Understanding Functions.
- Understanding Order of Operations
- Understanding Derivatives.
- Understanding Derivatives Notation.
- Understanding Common Derivatives; line, Quadratics, Higher Degree Polynomials, and Power Functions.

Number theory:

It goes all the way back to ancient times, when mathematicians studied different number systems such as:

- *Natural numbers*; These are the numbers 1, 2, 3, 4, 5...and so on. Only positive numbers are included.
- *Whole numbers*; Adding to natural numbers, the concept of “0”, for place-holding notation for empty “columns” on numbers greater than 9, such as “10,” “1,000,” or “1,090.” Those zeros indicate no value occupying that column.

- *Integers*; They include positive and negative natural numbers as well as 0.
- *Rational numbers*; Any number that can be expressed as a fraction, such as $2/3$, They are called rational because they are ratios.
- *Irrational numbers*; They cannot be expressed as a fraction. This includes π , square roots of certain numbers like $\sqrt{2}$, and Euler's number e , These numbers have an infinite number of decimal digits, such as 3.141592653589793238462....

- *Real numbers;* They include rational as well as irrational numbers. In data science treating any decimals as real numbers.
- *Complex and imaginary numbers;* Taking the square root of a negative number is an **imaginary number** ($\sqrt{-1}$), while The **complex number** is basically the combination of a real number and an imaginary number. The complex number is in the form of **$a + ib$** , where **a** and **b** are real number and **i** is imaginary number.

$$\underbrace{a}_{\text{Real part}} + \underbrace{ib}_{\text{Imaginary part}}$$

Order of Operations:

The order in which we solve math equations, there are six main operations that can be performed on numbers. These are; Addition, Subtraction, Multiplication, Division, Exponents, and Parentheses.

The order of operations can be easily recalled using the acronym **PEMDAS**. This stands for: **P**arentheses, **E**xponents, **M**ultiplication, **D**ivision, **A**ddition, and finally **S**ubtraction.

*Help you to remember **PEMDAS**: Please Excuse My Dear Aunt Sally*

P	()	PARENTHESES
E	x^2	EXPONENTS
M	$\times$	MULTIPLY
D	$\div$	DIVIDE
A	+	ADD
S	-	SUBTRACT

Examples

$$2 \times \frac{(3 + 2)^2}{5} - 4$$

1- Evaluate the Parentheses

$$2 \times \frac{(5)^2}{5} - 4$$

2- Solve the Exponents

$$2 \times \frac{25}{5} - 4$$

3- Perform Multiplication

$$\frac{50}{5} - 4$$

4- Perform Division

$$10 - 4 = 6$$

5- Perform Subtraction

P	()	PARENTHESES
E	x^2	EXPONENTS
M	$\times$	MULTIPLY
D	$\div$	DIVIDE
A	+	ADD
S	-	SUBTRACT

What is a variable?:

The variable is a named placeholder for an unspecified or unknown number.

Example

In a computer program, the program asks the user to enter an even number then the program will take this number and divide it by two.

Define **X** as an **Integer**

```
X= input("enter an even number:")
```

```
Result= X / 2
```

X and **Result** are variables

What is a function?:

The function is an expression that defines the relationship between one or more variables.

Example

For linear equation : $y = 2x + 1$

For any given x value, we can solve the equation and find the value of y

The value of y depends on the value of x , then y is a function of x .

Thus, we can rewrite the linear equation to be:

$$f(x) = 2x + 1$$

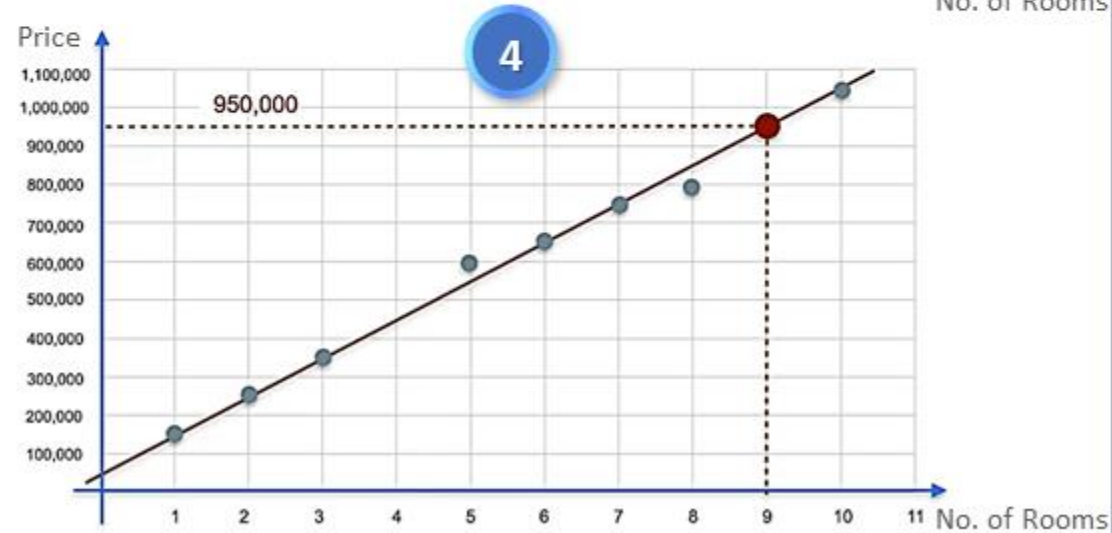
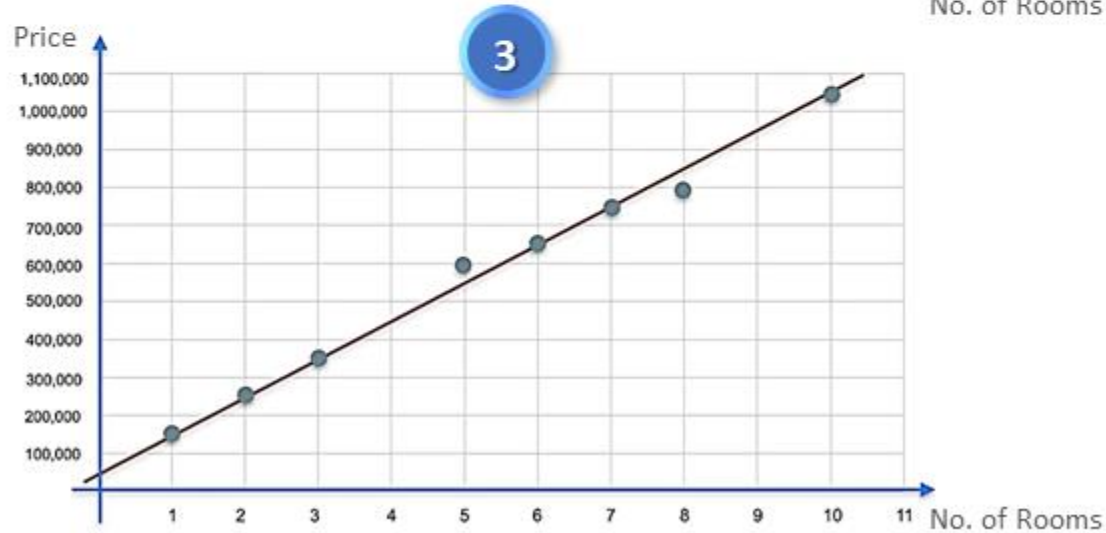
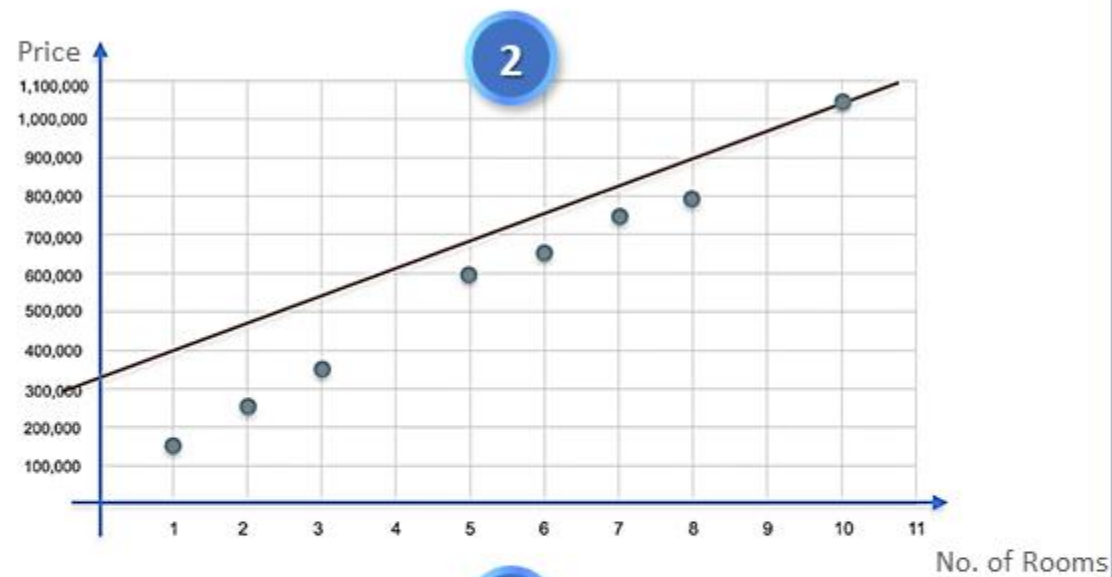
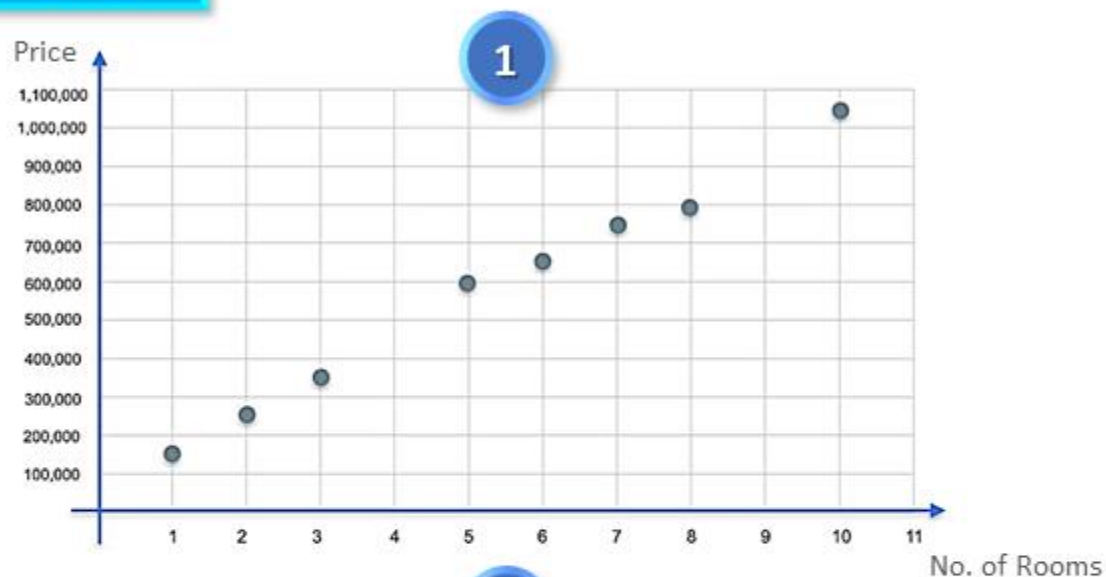
e.g., $f(5) = 2 \times 5 + 1 = 11$

x	expression	y
0	$2 \times 0 + 1$	1
1	$2 \times 1 + 1$	3
2	$2 \times 2 + 1$	5
3	$2 \times 3 + 1$	7

Why Derivatives is important to data science?:

Derivatives are used to optimize functions to maximize them and minimize them. That means *finding the maximum value or the minimum value of a function*. This is very important in data science.

Introduction to Derivatives



The question is, at a given instant, what is the velocity at that instant? That's called the instantaneous velocity. That is precisely what a derivative is. ***A derivative is the instantaneous rate of change of a function.*** In this case, the function is distance and its derivative is velocity.

*Time*

$$\text{Speed} = \frac{\text{Distance}}{\text{Time}}$$



Introduction to Derivatives

t (seconds)	x (meters)
0	0
5	36
10	122
15	202
20	265
25	351
30	441
35	551
40	591
45	716
50	816
55	900
60	1000

$$10-15 \text{ sec} = 202-122=80$$

$$15-20 \text{ sec} = 265-202=63$$



Is the car moving at constant speed?

t (seconds)	x (meters)
0	0
5	36
10	122
15	202
20	265
25	351
30	441
35	551
40	591
45	716
50	816
55	900
60	1000

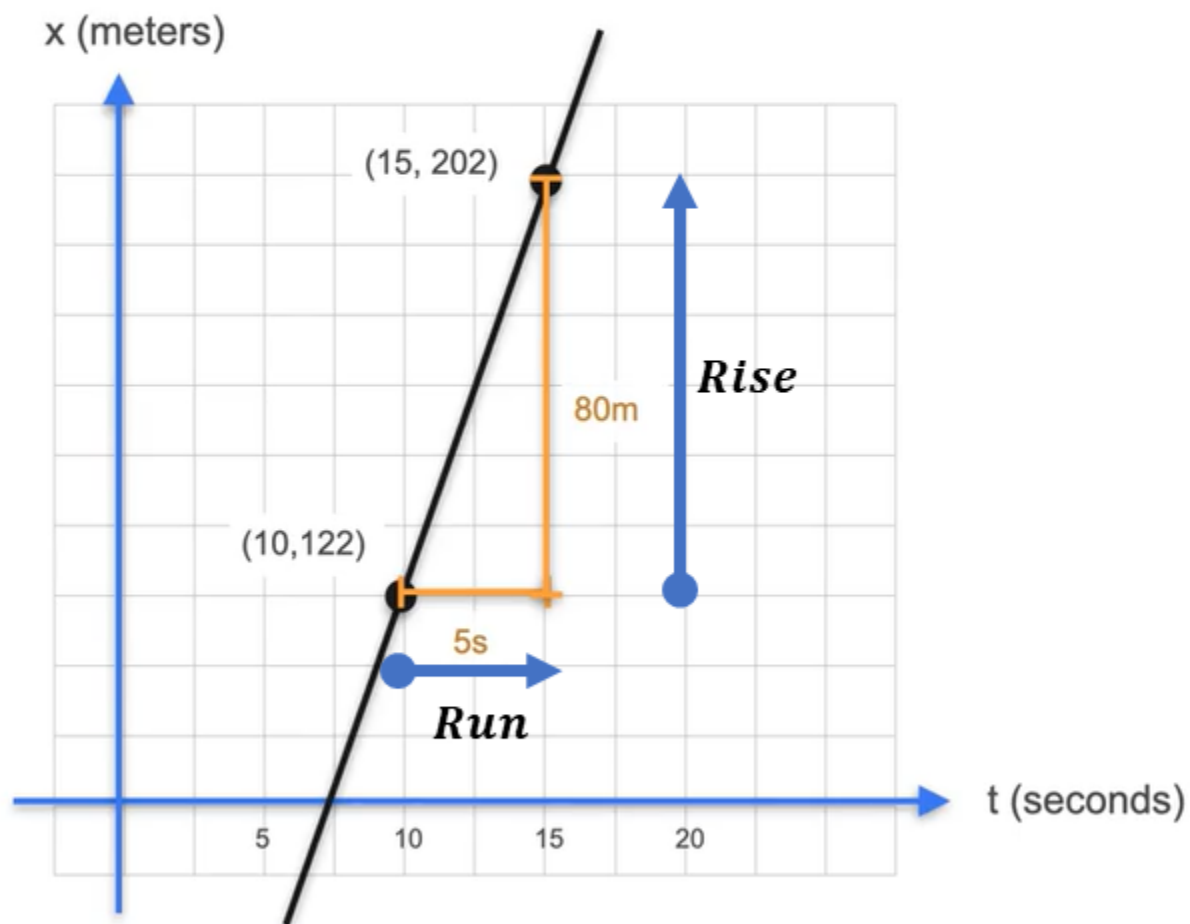
Can we use the Information presented to calculate the velocity at 12.5 sec?

$$\text{Speed} = \frac{\text{Distance}}{\text{Time}}$$

No, we can not, we may calculate the average velocity between two-time intervals but not at the specific exact time.

For example, we can calculate the average velocity between 10 – 15 seconds.

Introduction to Derivatives



$$\text{Speed} = \frac{\text{Distance}}{\text{Time}} = \text{Slope} = \frac{\text{Rise} \uparrow}{\text{Run} \rightarrow}$$

$$\text{Slope} = \frac{\text{Change in distance } (\Delta x)}{\text{Change in Time } (\Delta t)}$$

$$\text{Slope} = \frac{x(15) - x(10)}{t(15) - t(10)}$$

$$\text{Slope} = \frac{202\text{ m} - 122\text{ m}}{15\text{ sec.} - 10\text{ sec.}}$$

$$\text{Slope} = \frac{80\text{ m}}{5\text{ sec.}} = 16\text{ m/Sec.}$$

Can we find a better estimate of velocity at 12.5 Seconds?

t (s)	x(m)
10	122
11	138
12	155
13	170
14	186
15	202
16	218
17	234
18	250
19	265
20	265

$$\text{Slope} = \frac{\text{Change in distance } (\Delta x)}{\text{Change in Time } (\Delta t)}$$

$$\text{Slope} = \frac{x(13) - x(12)}{t(13) - t(12)}$$

$$\text{Slope} = \frac{170\text{m} - 155\text{m}}{13\text{sec.} - 12\text{sec.}}$$

$$\text{Slope} = \frac{15\text{m}}{1\text{sec.}} = 15 \text{ m/Sec.}$$

We still do not have the estimate of velocity at 12.5 Seconds.

Just take finer and finer intervals and the finer the intervals, the better the estimate is.

Slopes, Maxima, and Minima

t (s)	x (m)
19	265
20	265

Between 19 seconds and 20 seconds, the distance at 19 is 265 and the distance of 20 is 265. The distance is the same at both points.

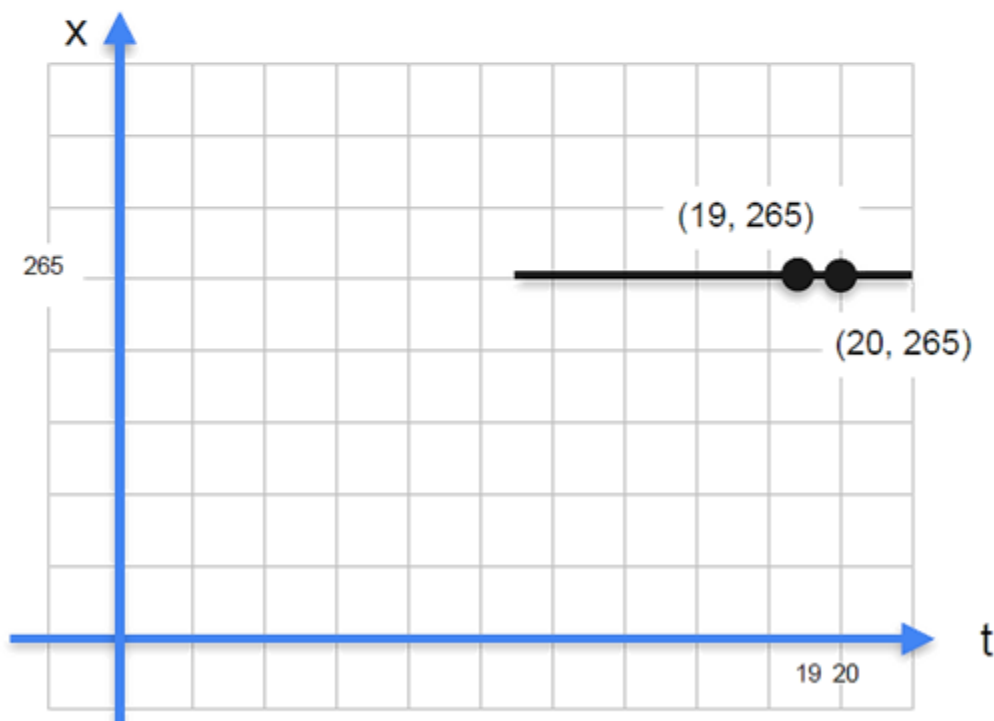
$$\text{Slope} = \frac{\Delta x}{\Delta t}$$

$$\text{Slope} = \frac{x(20) - x(19)}{20 - 19}$$

$$\text{Slope} = \frac{265 - 265}{1}$$

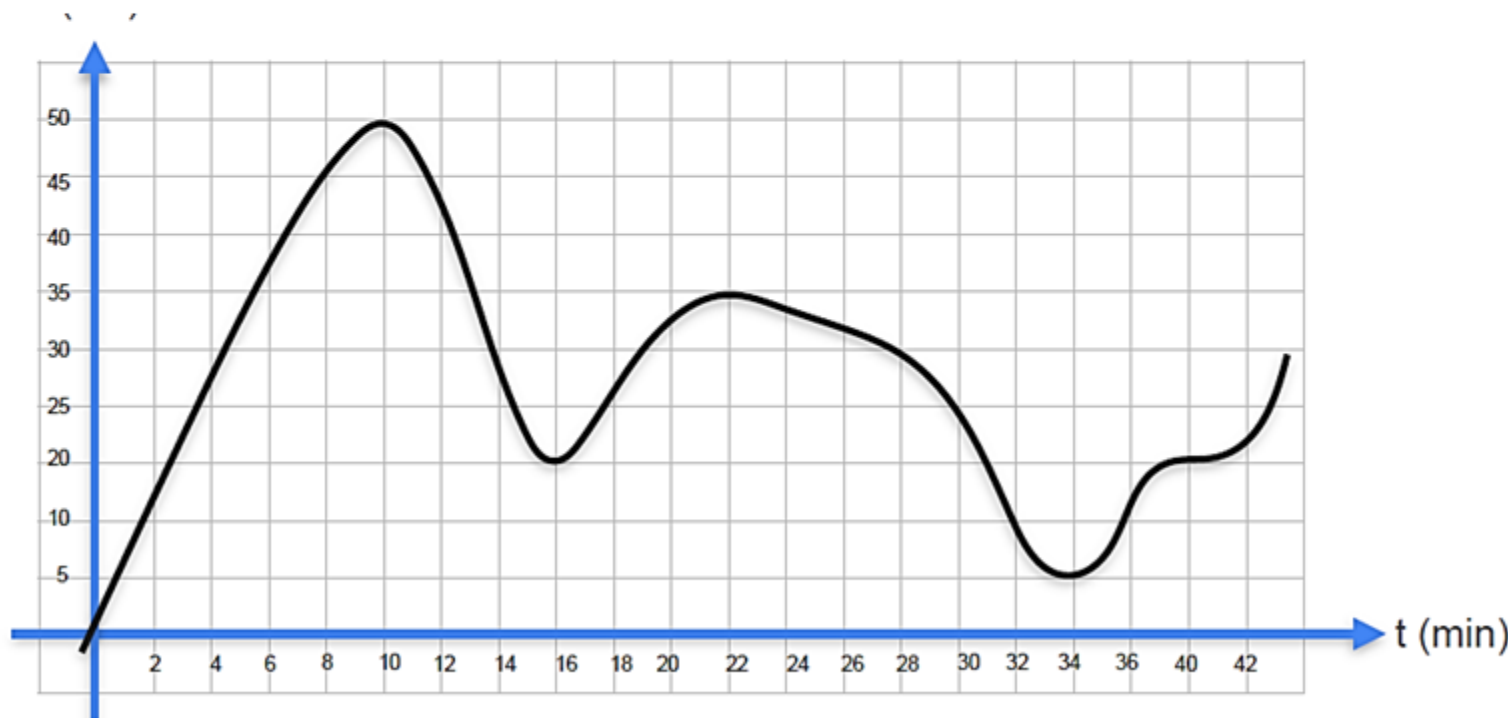
$$\text{Slope} = \frac{0}{1} = 0$$

$$\text{Slope} = 0 \text{ m/s}$$



Slopes, Maxima, and Minima

First the car go forward, then it stops, then it goes backwards, and then it stops, goes forward again, and then backwards again and then forwards again. Here is the question, *where was the velocity of the car zero?*

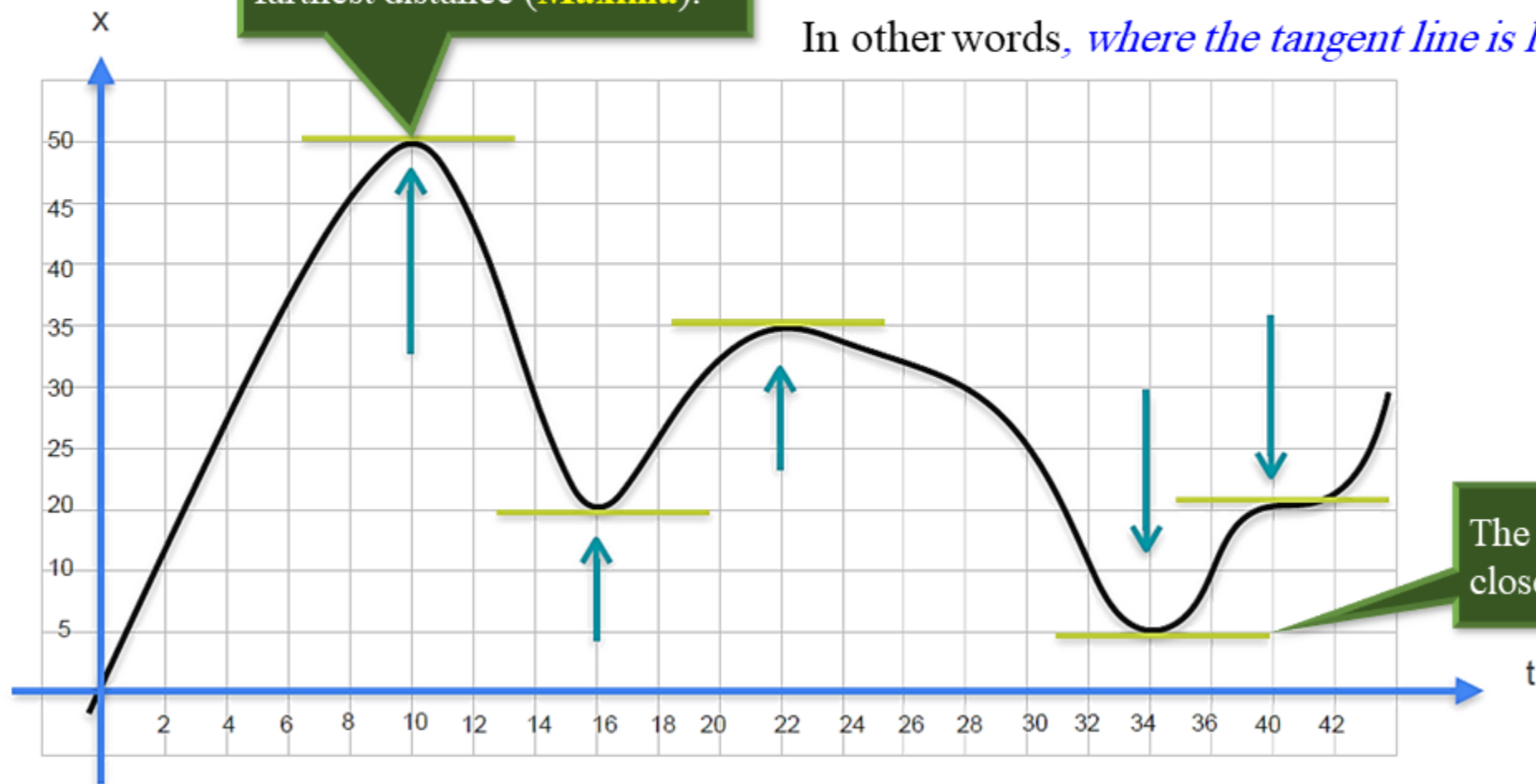


Slopes, Maxima, and Minima

The time where the car at the farthest distance (**Maxima**).

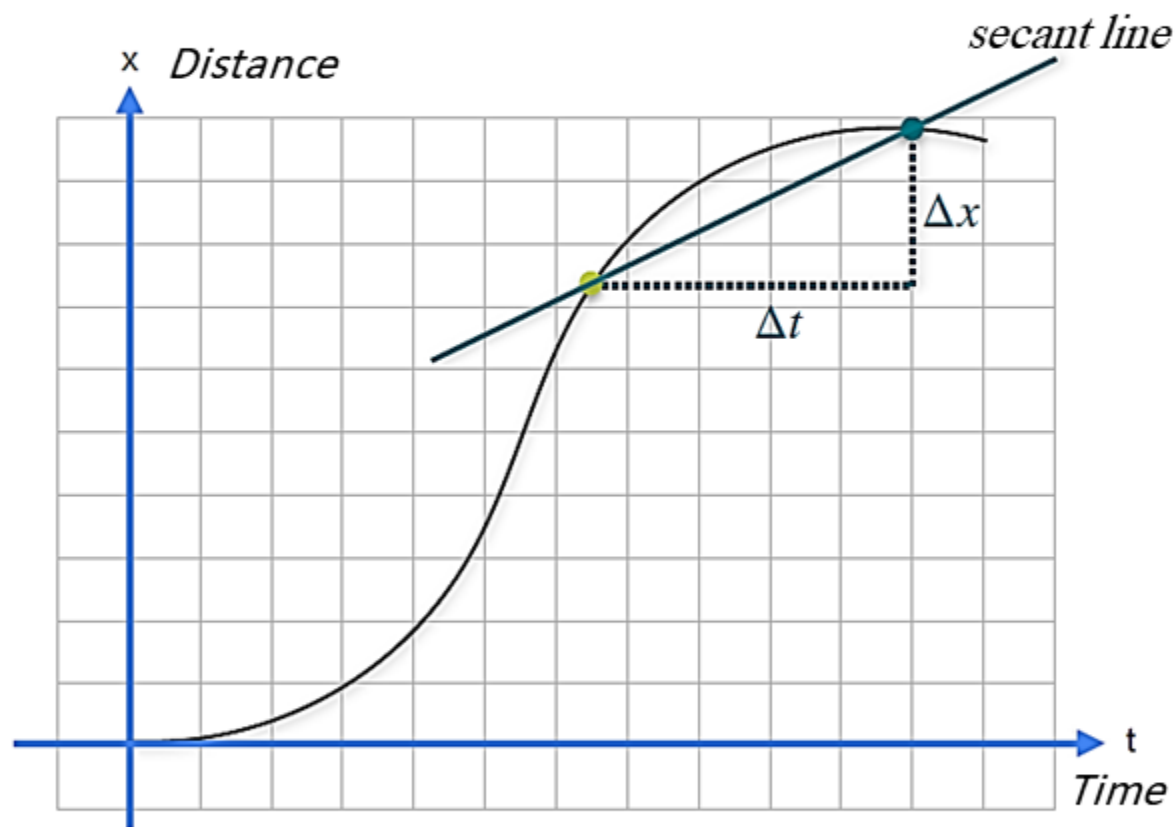
Finding the maximum or the minimum in a function, *it occurs at one of the points where the derivative is zero.*

In other words, *where the tangent line is horizontal.*



The time where the car at the closest distance (**Minima**).

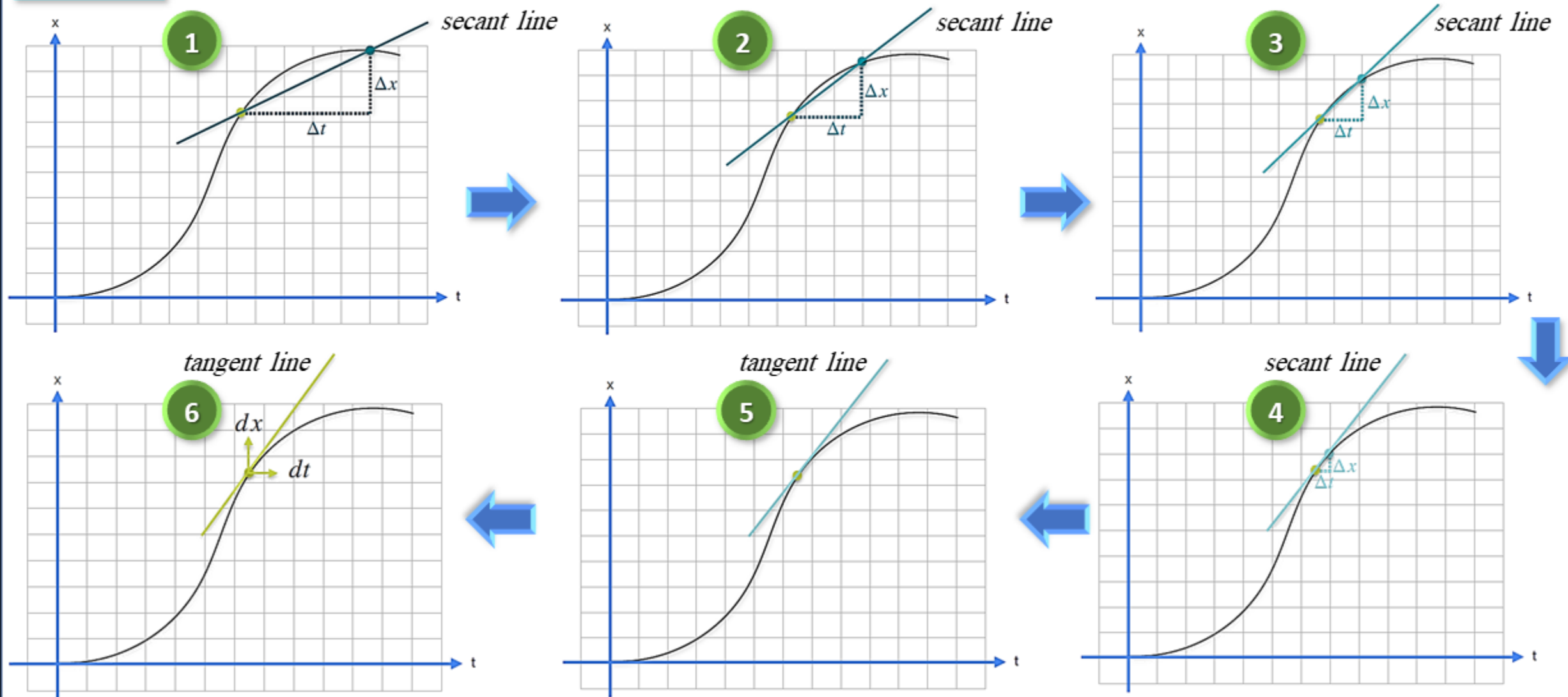
Derivatives Notation

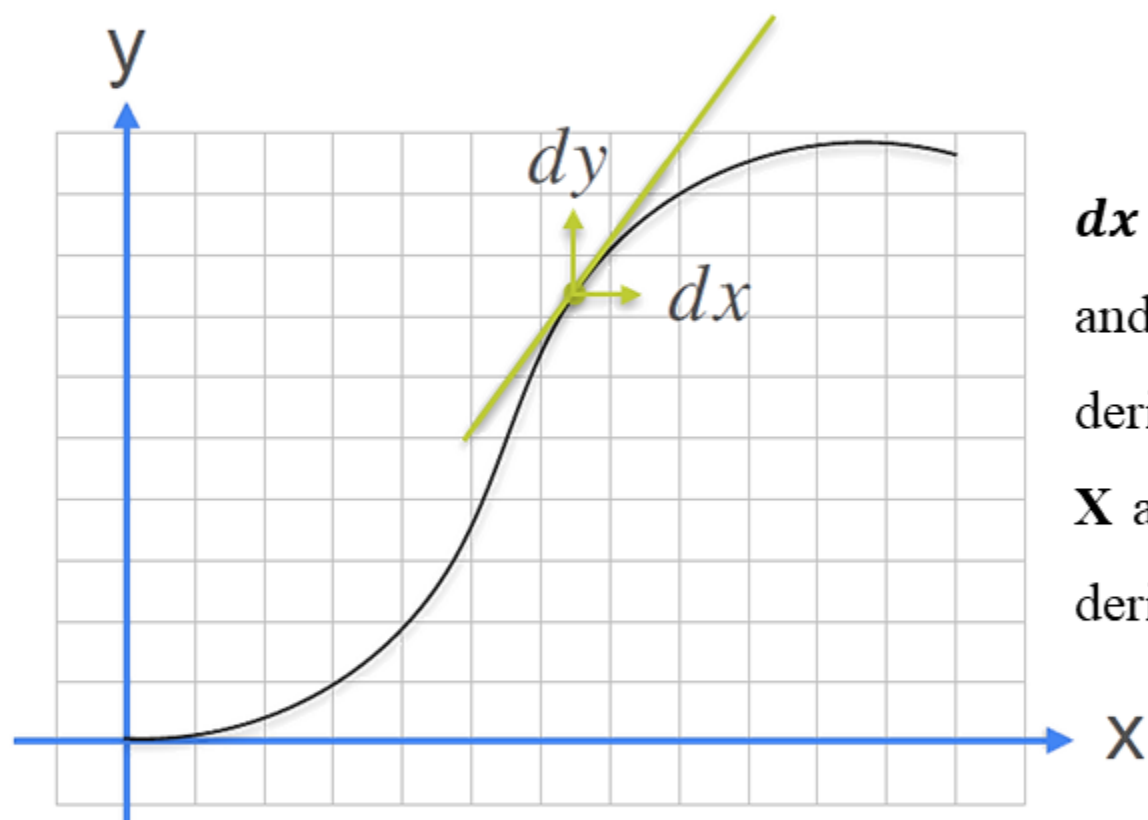


$$\text{Slope} = \frac{\text{Change in distance } (\Delta x)}{\text{Change in Time } (\Delta t)}$$

In order to calculate the slope, we first start with a big interval calculated Δx over Δt , that's the slope of that secant (*A secant line is a line through two points on the curve*). And then made the interval smaller and smaller, always calculating Δx over Δt . Until the interval was so tiny that you get the tangent line (*A tangent line is a line that touches a curve at only one point*), and this slope was called dx over dt .

Derivatives Notation



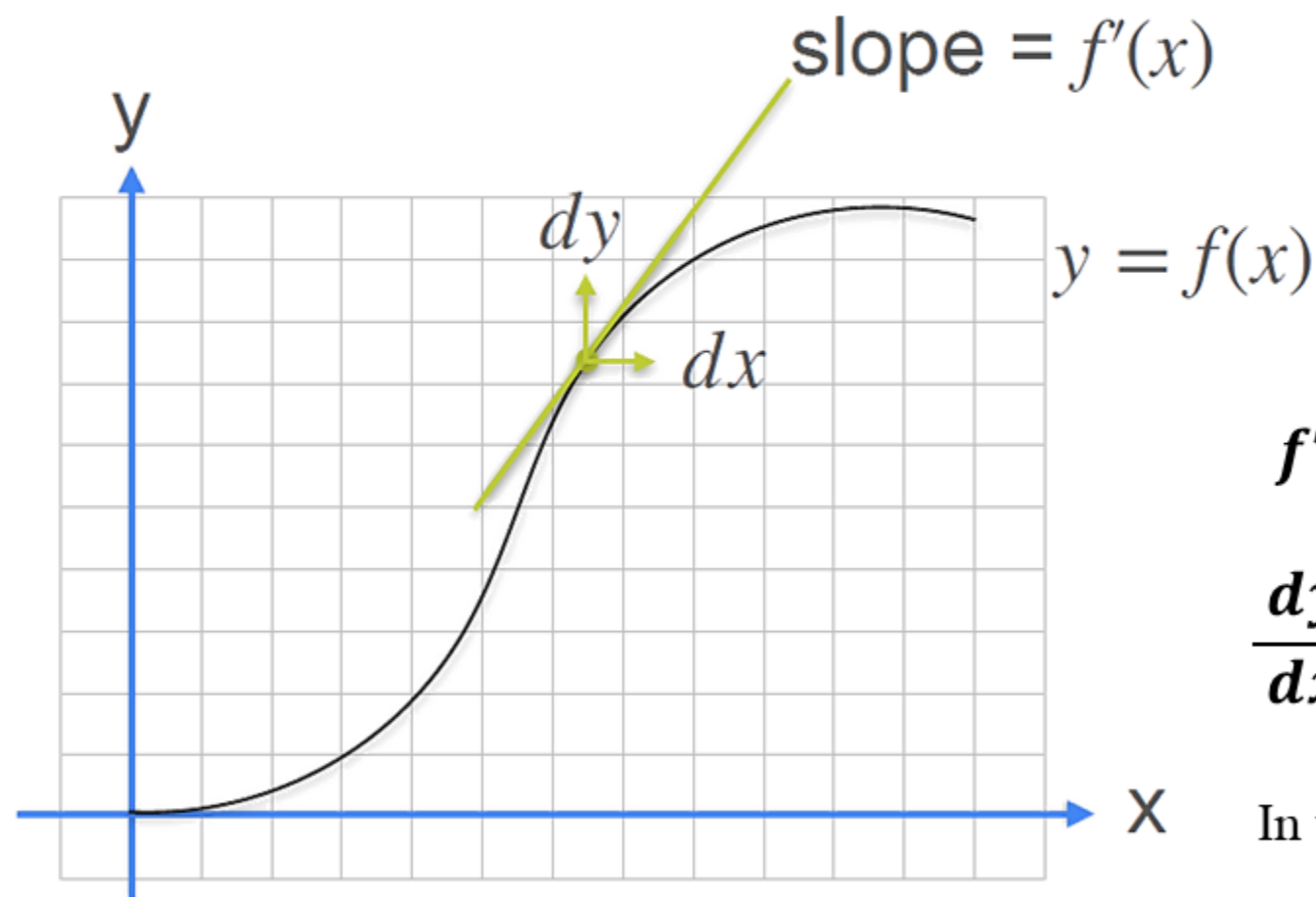


dx over dt represents infinitesimal changes in the **x-direction** and in the **t-direction**. So that's one way of writing the derivative as dx over dt . Now in general we're going to have **X** as the horizontal axis and **Y** as the vertical axis. So, the derivative is going to be called dy over dx instead.

$$\text{Slope at point} = \frac{\text{vertical change}}{\text{horizontal change}}$$

$$\text{Slope at point} = \frac{\Delta y}{\Delta x} = \frac{dy}{dx}$$

Derivatives Notation



$$y = f(x)$$

Derivative of f is expressed as

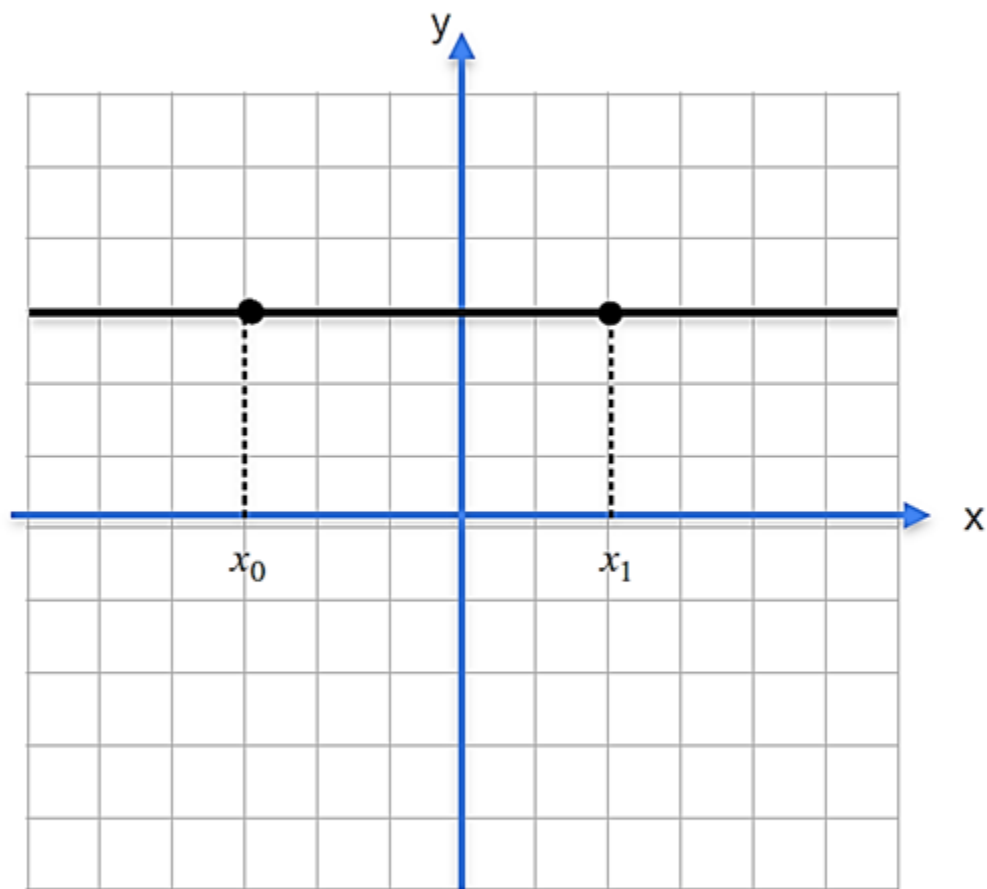
$$f'(x)$$

Lagrange's notation

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

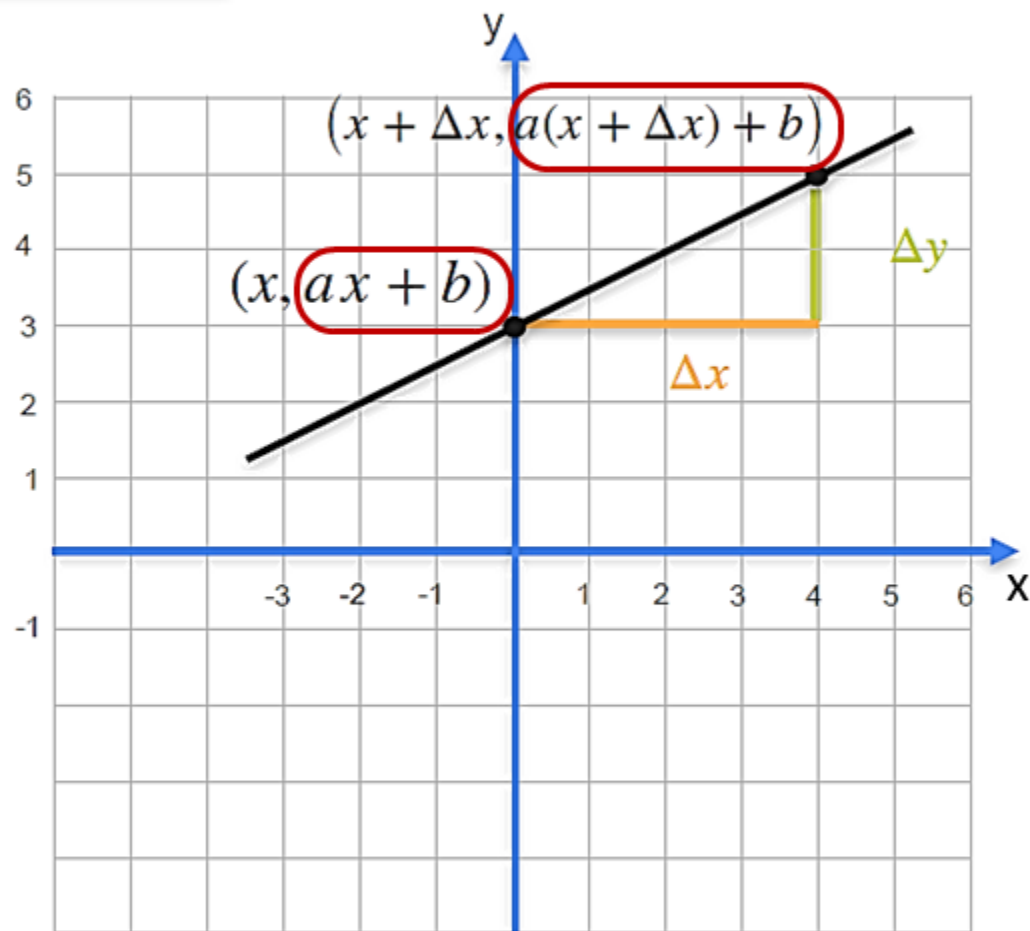
Leibniz's notation

In this case $\frac{d}{dx}$ can be thought of as an operator



In this case, y is constant, so $y = f(x) = c$, thus:

$$\frac{dy}{dx} = \frac{c - c}{x_1 - x_0} = \frac{0}{\Delta x} = 0 \rightarrow f'(x) = 0$$



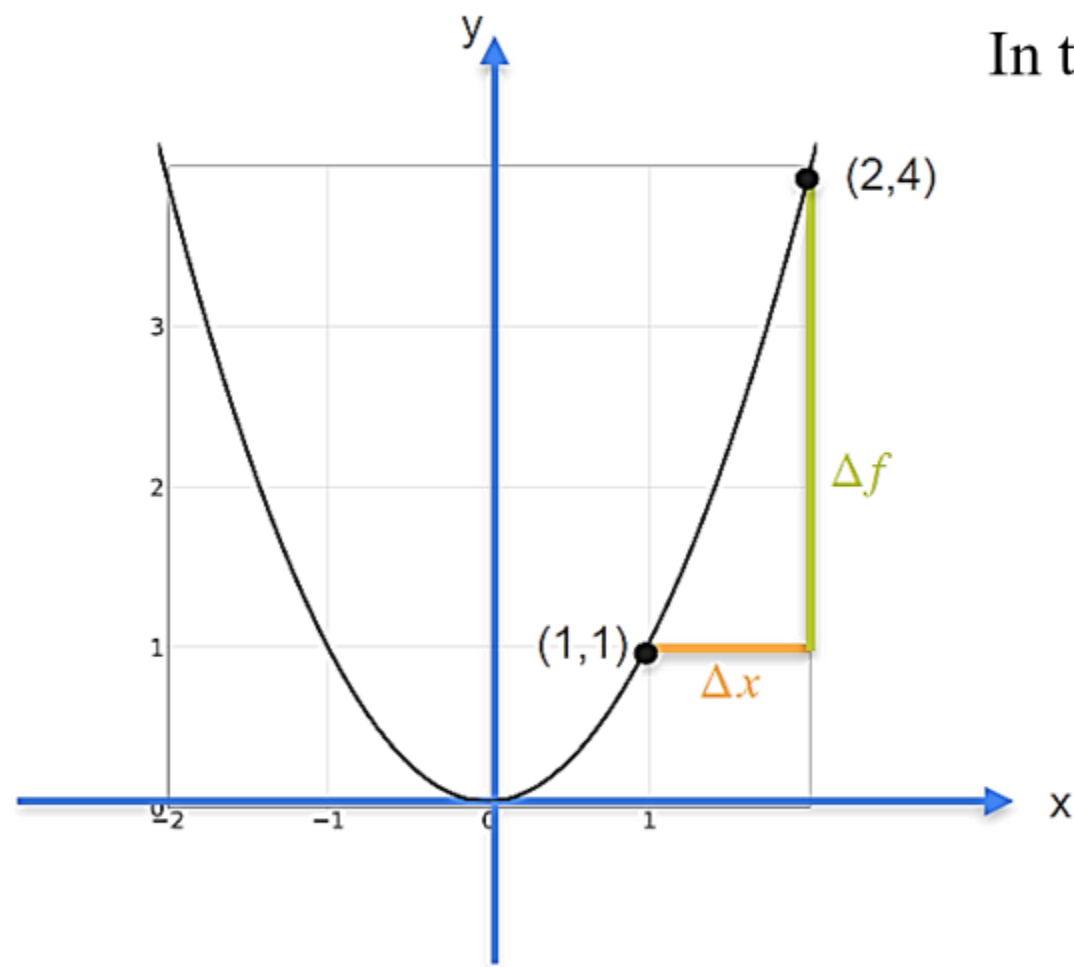
In this case, $f(x) = ax + b$, so $f'(x) = a$, thus:

$$\frac{\Delta y}{\Delta x} = \frac{\text{Rise}}{\text{Run}} = a$$

$$\frac{\Delta y}{\Delta x} = \frac{\cancel{ax} + a\Delta x + \cancel{b} - \cancel{ax} - \cancel{b}}{\Delta x}$$

$$\frac{\Delta y}{\Delta x} = \frac{a\cancel{\Delta x}}{\cancel{\Delta x}} = a$$

Common Derivatives; Quadratics



In this case, $y = f(x) = x^2$, so $f'(x) = 2x$, thus:

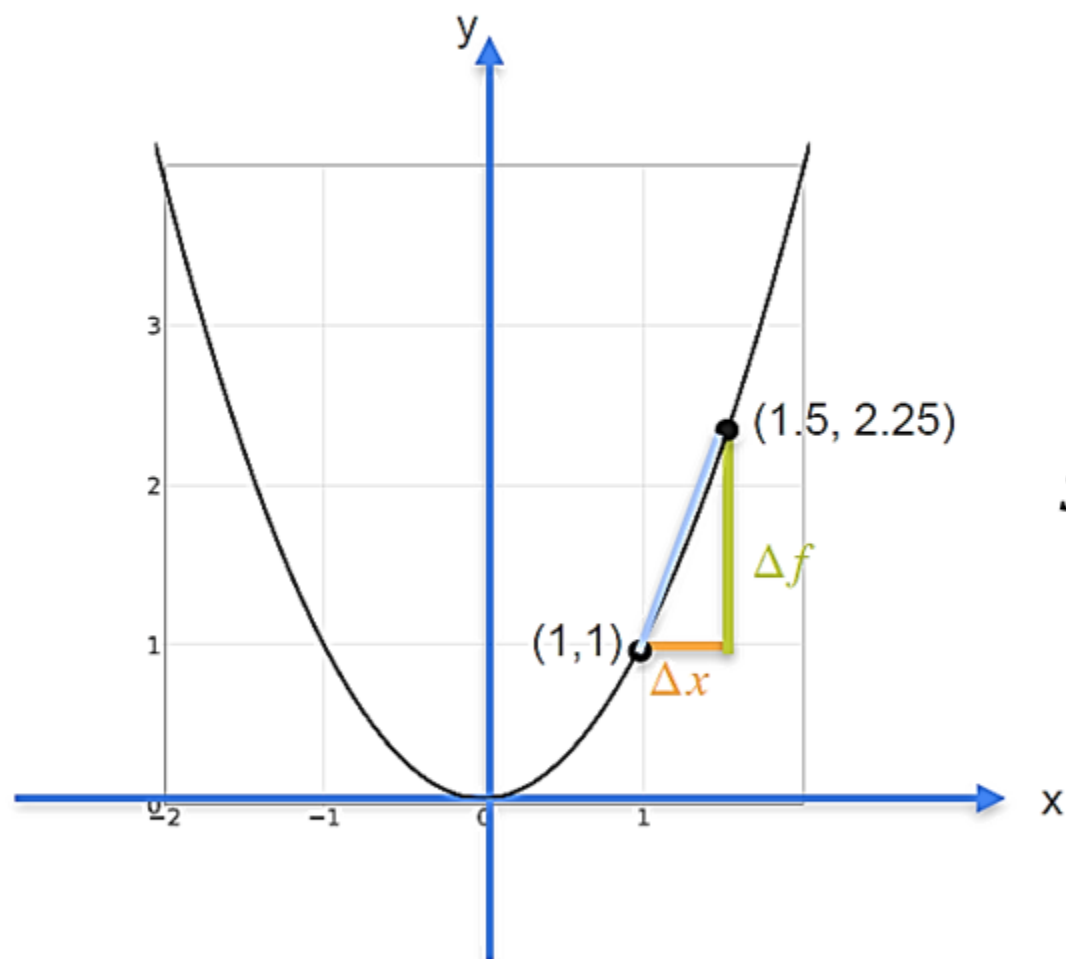
$$\text{Slope} = \frac{df}{dx} = \frac{(x + \Delta x)^2 - (x)^2}{\Delta x}$$

$$df = (1 + 1)^2 - (1)^2 = 4 - 1 = 3$$

$$\text{Slope} = \frac{3}{1} = 3$$

Δx	1
Δf	3
Slope	3

Common Derivatives; Quadratics

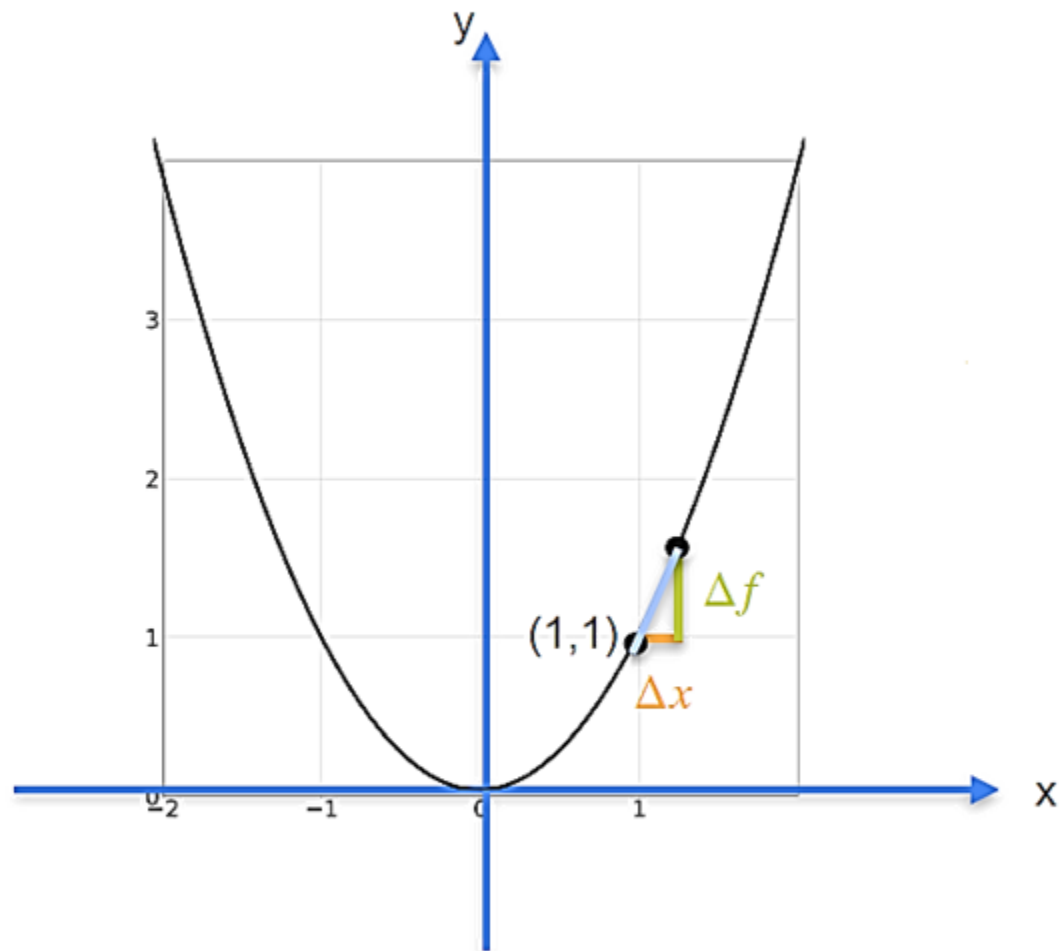


$$\text{Slope} = \frac{df}{dx} = \frac{(x + \Delta x)^2 - (x)^2}{\Delta x}$$

$$df = (1 + 0.5)^2 - (1)^2 = 2.25 - 1 = 1.25$$

$$\text{Slope} = \frac{1.25}{0.5} = 2.5$$

Δx	1	$1/2$
Δf	3	1.25
Slope	3	2.5



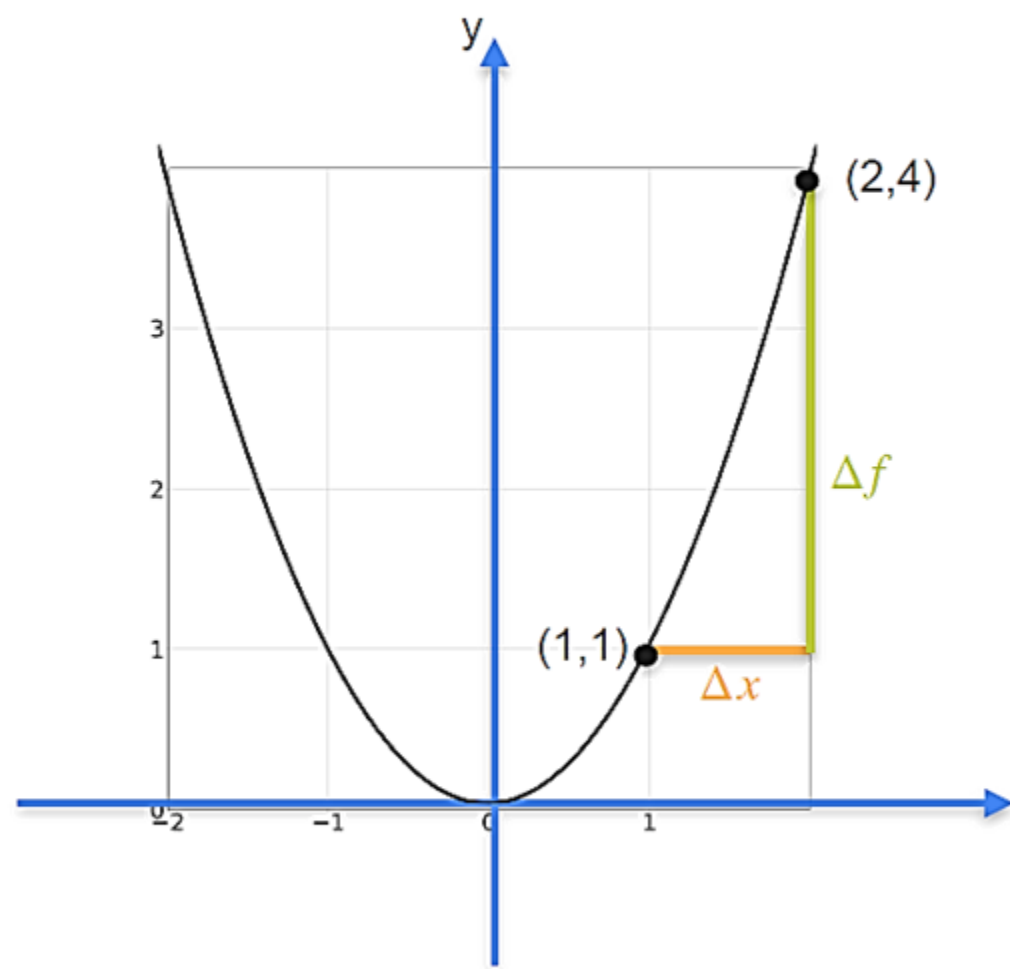
$$\text{Slope} = \frac{df}{dx} = \frac{(x + \Delta x)^2 - (x)^2}{\Delta x}$$

Δx	1	$1/2$	$1/4$	$1/8$	$1/16$	$1/1000$
Δf	3	1.25	0.562	0.265	0.128	0.002
Slope	3	2.5	2.25	2.125	2.065	2.001

$$f'(x) = \frac{d}{dx} f(x) = 2x$$

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

Common Derivatives; Quadratics



In this case, $y = f(x) = x^2$

$$\frac{df}{dx} = \frac{(x + \Delta x)^2 - (x)^2}{\Delta x}$$

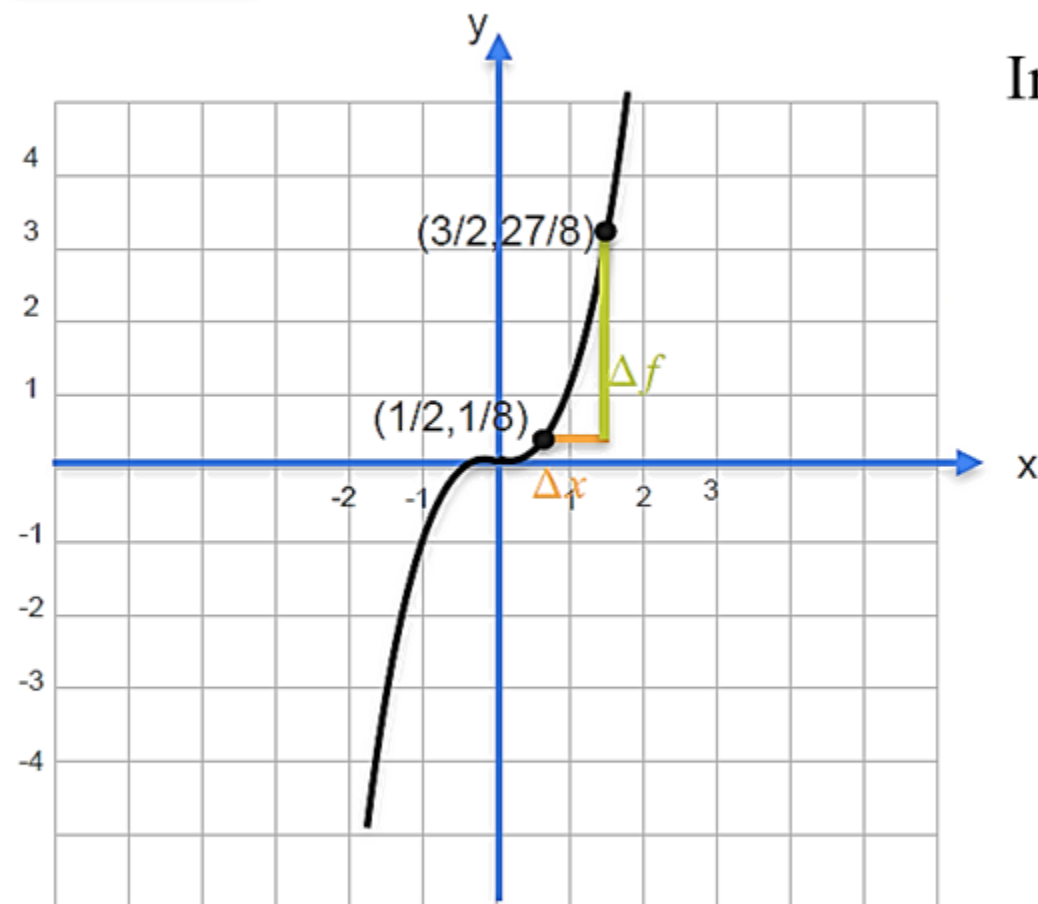
$$\begin{aligned} \frac{df}{dx} &= \frac{\cancel{x^2} + 2x\cancel{\Delta x} + (\Delta x)^2 - \cancel{x^2}}{\cancel{\Delta x}} \\ &= 2x + \Delta x \end{aligned}$$

As Δx goes to 0, the $2x$ stays put and this Δx is 0.

Therefore, the derivative $\frac{df}{dx}$ is going to be $2x$.

$$f(x) = x^2 \longrightarrow f'(x) = 2x$$

Common Derivatives; Higher Degree Polynomials



In this case, $y = f(x) = x^3$, so $f'(x) = 3x^2$, thus:

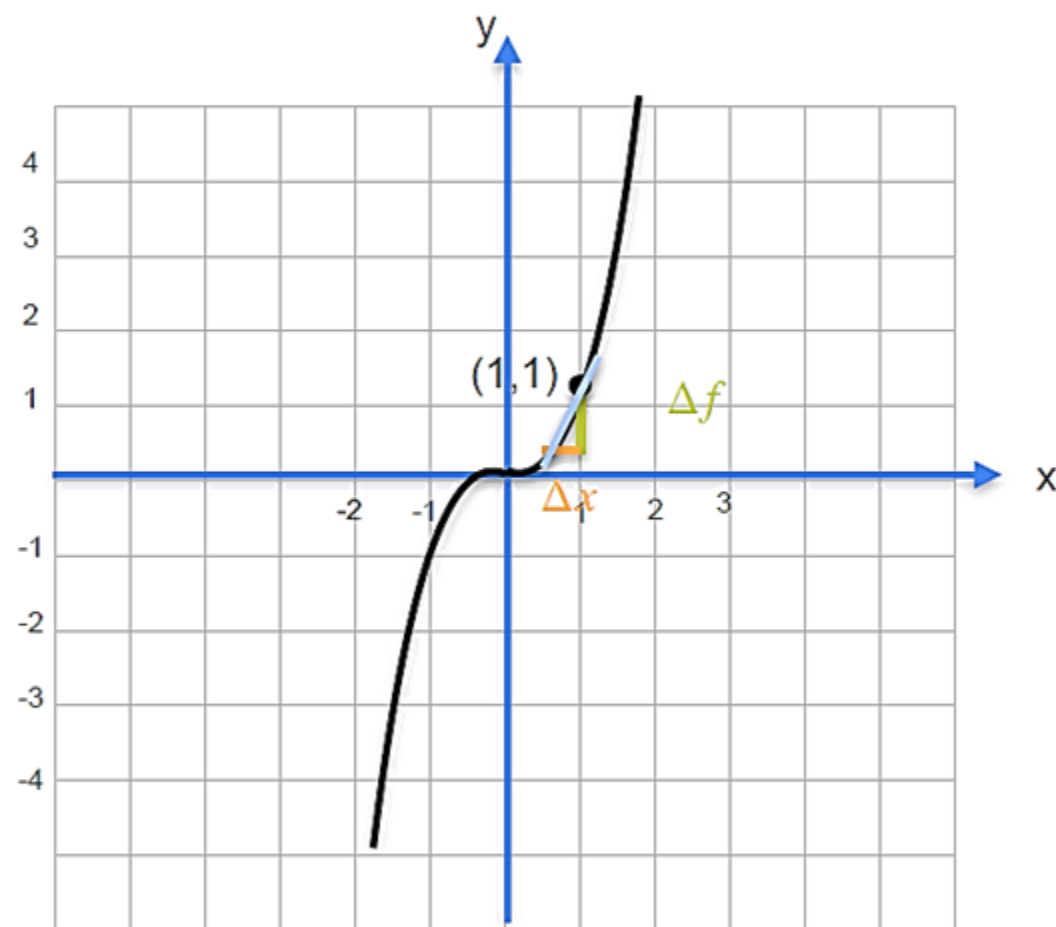
$$\text{Slope} = \frac{df}{dx} = \frac{(x + \Delta x)^3 - (x)^3}{\Delta x}$$

$$df = \left(\frac{1}{2} + 1\right)^3 - \left(\frac{1}{2}\right)^3 = \frac{27}{8} - \frac{1}{8} = 3.25$$

$$\text{Slope} = \frac{3.25}{1} = 3.25$$

Δx	1
Δf	3.25
Slope	3.25

Common Derivatives; Higher Degree Polynomials



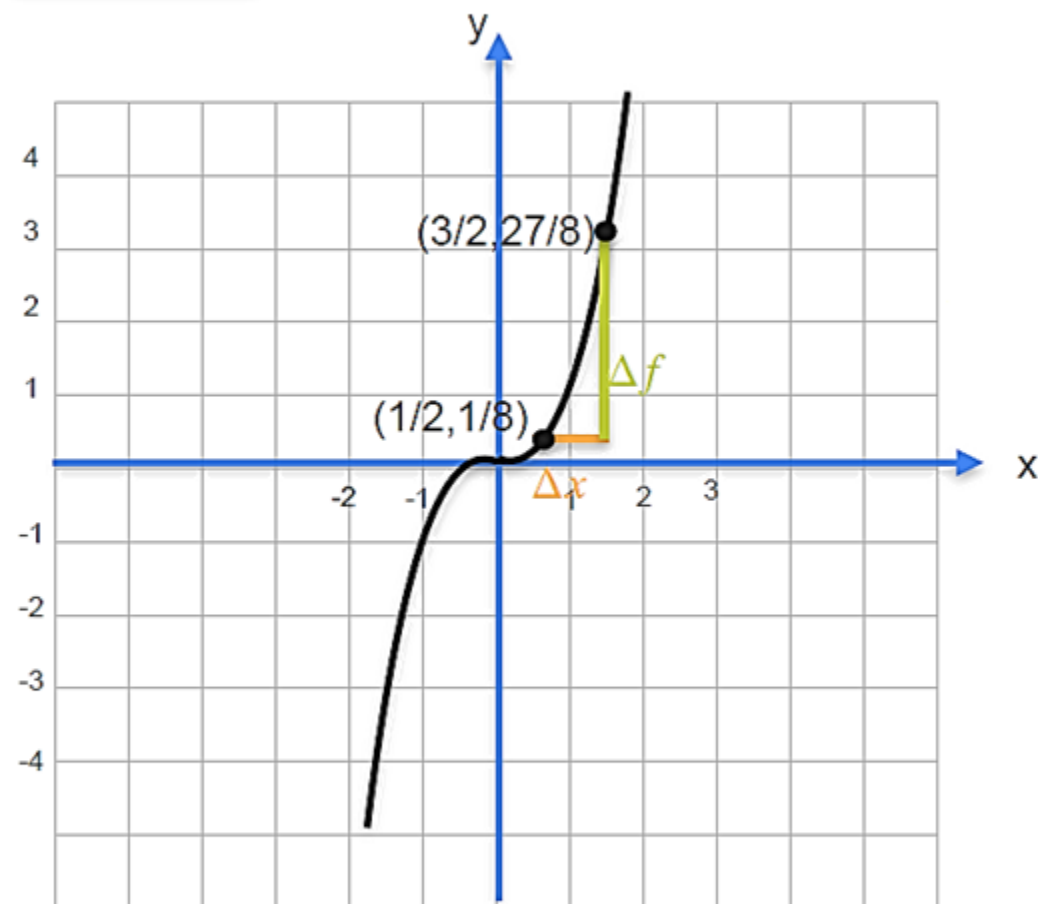
$$\text{Slope} = \frac{df}{dx} = \frac{(x + \Delta x)^3 - (x)^3}{\Delta x}$$

Δx	1	$1/2$	$1/4$	$1/8$	$1/16$	$1/1000$
Δf	3.25	0.86	0.30	0.12	0.05	0.0008
Slope	3.25	1.75	1.188	0.95	0.85	0.75

$$f'(x) = \frac{d}{dx} f(x) = 3x^2$$

$$f'(0.5) = \frac{d}{dx} f(0.5) = 3 \times (0.5)^2 = 0.75$$

Common Derivatives; Higher Degree Polynomials



$$\text{Slope} = \frac{df}{dx} = \frac{(x + \Delta x)^3 - (x)^3}{\Delta x}$$

$$\frac{df}{dx} = \frac{\cancel{x^3} + 3x(\Delta x)^2 + 3x^2\cancel{\Delta x} + (\Delta x)^3 - \cancel{x^3}}{\Delta x}$$

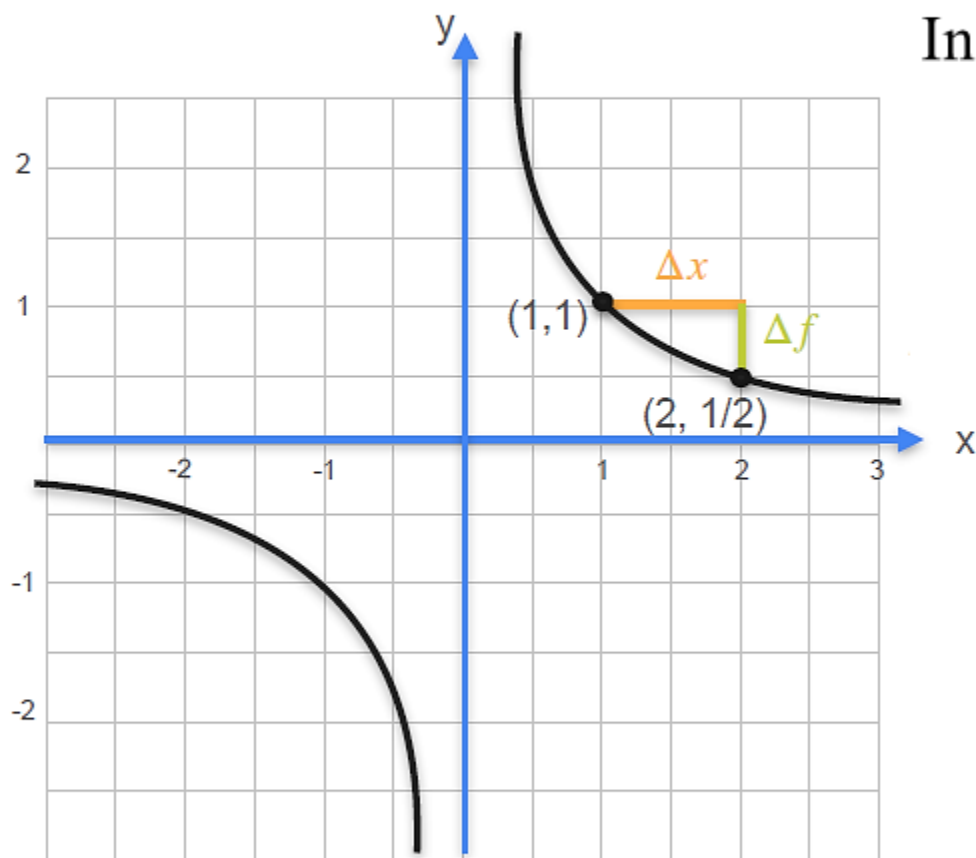
$$\frac{df}{dx} = 3x\Delta x + 3x^2 + (\Delta x)^2$$

as $\Delta x \rightarrow 0$

$$\frac{df}{dx} = 3x\cancel{\Delta x} + 3x^2 + (\cancel{\Delta x})^2$$

$$f'(x) = \frac{d}{dx} f(x) = 3x^2$$

Common Derivatives; Other Power Functions



In this case, $y = f(x) = \frac{1}{x} = x^{-1}$, so $f'(x) = -x^{-2}$, thus:

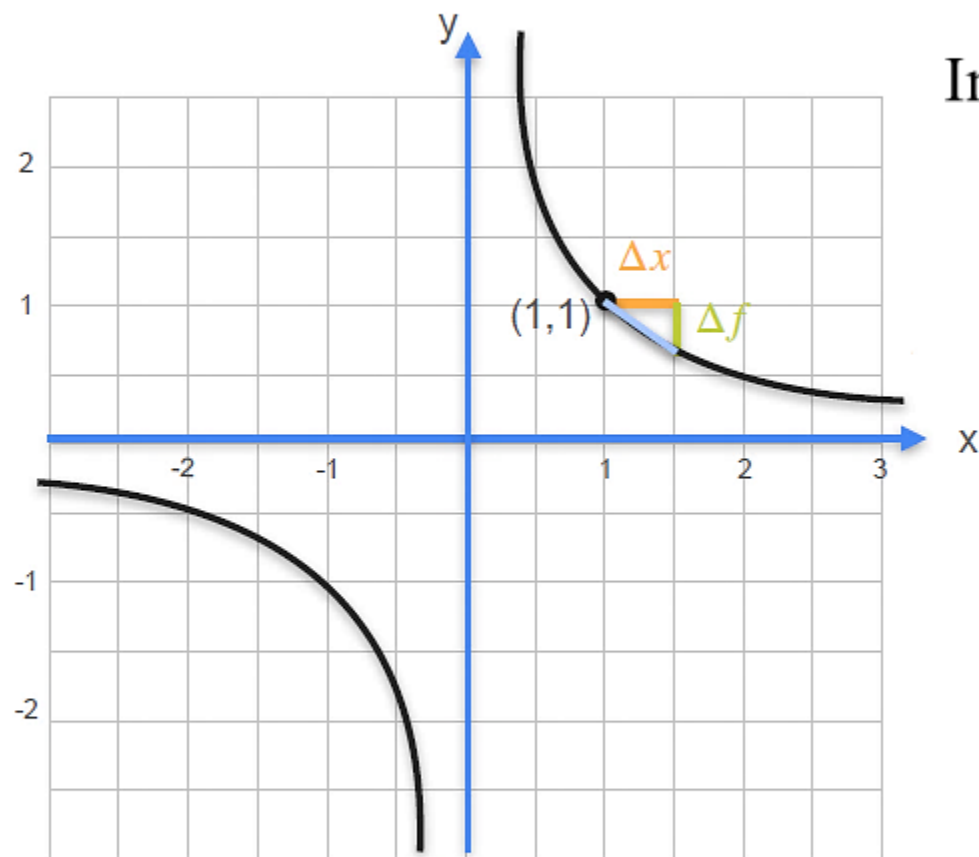
$$\text{Slope} = \frac{\Delta f}{\Delta x} = \frac{(x + \Delta x)^{-1} - (x)^{-1}}{\Delta x}$$

Δx	1
Δf	-0.5
<i>Slope</i>	-0.5

$$df = (1 + 1)^{-1} - 1^{-1} = \frac{1}{2} - 1 = -0.5$$

$$\text{Slope} = \frac{dy}{dx} = \frac{-0.5}{1} = -0.5$$

Common Derivatives; Other Power Functions



In this case, $y = f(x) = \frac{1}{x} = x^{-1}$, so $f'(x) = -x^{-2}$, thus:

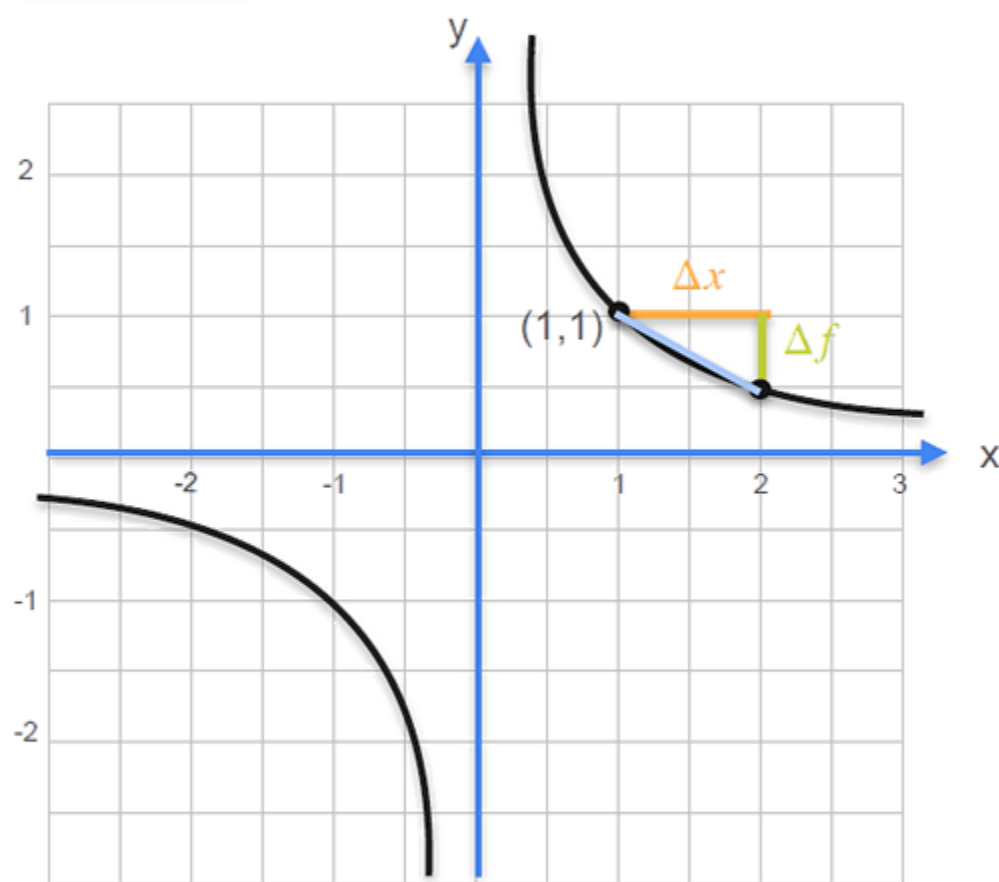
$$\text{Slope} = \frac{\Delta f}{\Delta x} = \frac{(x + \Delta x)^{-1} - (x)^{-1}}{\Delta x}$$

Δx	1	$1/2$	$1/4$	$1/8$	$1/16$	$1/1000$
Δf	-0.5	-0.33	-0.2	-0.11	-0.06	-0.001
Slope	-0.5	-0.67	-0.8	-0.89	-0.94	-0.999

$$df = (1 + 1)^{-1} - 1^{-1} = \frac{1}{2} - 1 = -0.5$$

$$\text{Slope} = \frac{dy}{dx} = \frac{-0.5}{1} = -0.5$$

Common Derivatives; Other Power Functions



$$\text{Slope} = \frac{\Delta f}{\Delta x} = \frac{(x + \Delta x)^{-1} - (x)^{-1}}{\Delta x}$$

$$\begin{aligned} \text{Slope} = \frac{\Delta f}{\Delta x} &= \frac{\frac{1}{x + \Delta x} - \frac{1}{x}}{\Delta x} = \frac{\cancel{x} - (\cancel{x} + \Delta x)}{(x + \Delta x)x \Delta x} \\ &= -\frac{1}{x^2 + x\Delta x} \end{aligned}$$

as $\Delta x \rightarrow 0$

$$= -\frac{1}{x^2}$$

$$f(x) = \frac{1}{x} = x^{-1} \Rightarrow f'(x) = -\frac{1}{x^2} = -x^{-2}$$

Derivatives of Power Functions

$$f(x) = x^2$$

$$f'(x) = 2x^1$$

$$f(x) = x^3$$

$$f'(x) = 3x^2$$

$$f(x) = x^{-1}$$

$$f'(x) = (-1)x^{-2}$$

General derivative of power function:

$$f(x) = x^n \quad \longrightarrow \quad f'(x) = \frac{d}{dx} f(x) = nx^{n-1}$$

Exercises

What are the derivatives of the following functions?

$$f(x) = 2x^2$$

$$f'(x) = \dots \dots$$

$$f(x) = -x^3$$

$$f'(x) = \dots \dots$$

$$f(x) = 10x^2$$

$$f'(x) = \dots \dots$$

$$f(x) = -x^{-2}$$

$$f'(x) = \dots \dots$$

