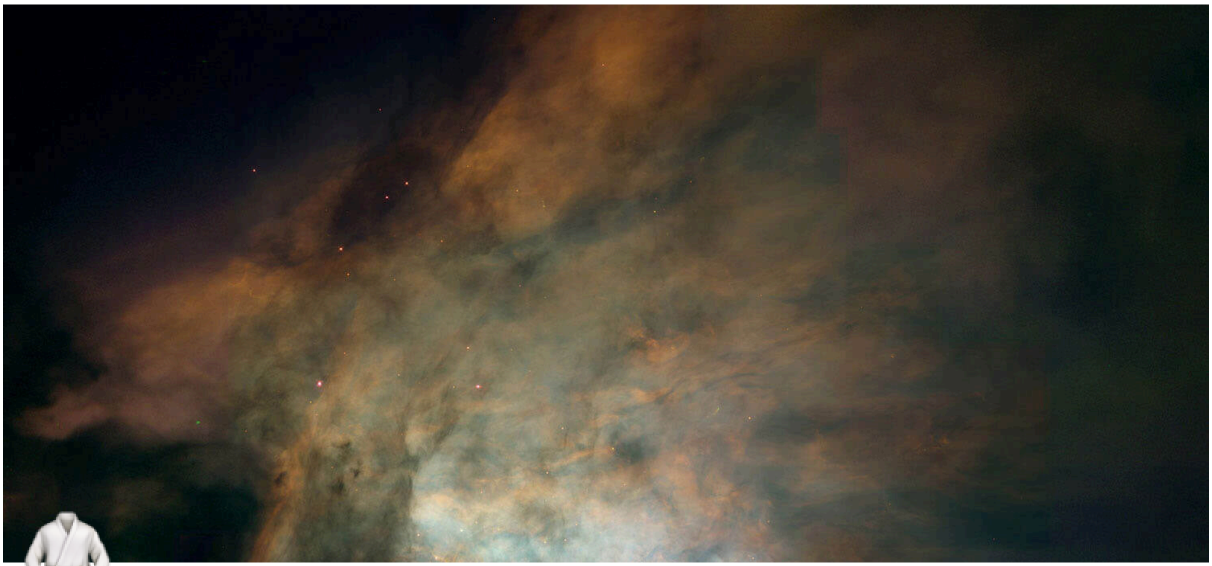
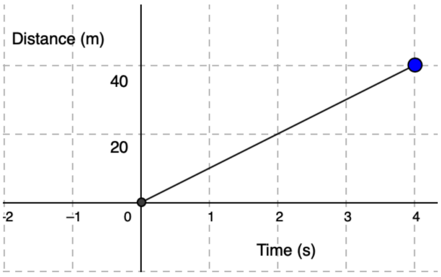


Introduction to differentiation

Rate of Change (Gradient) of a Straight Line

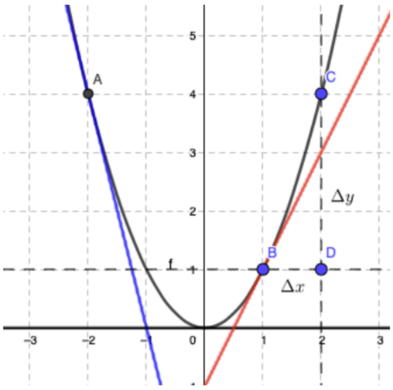
- *A car travels 40m over 4s. Speed?*
- *Gradient (derivative)= speed/slope = rate of distance change, steepness*
- Δx, Δy: *change of x and y*
- *Gradient of a straight line: y = 10x*
    - *Constant gradient, same at every point.*



UNIVERSITY OF BIRMINGHAM

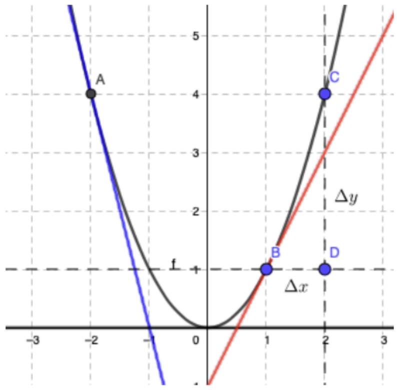
Gradient at a Point, Differentiation from first principles

- The gradient at a point is given by the gradient of the tangent at that point.
- As point C moves closer to B, the gradient of the line BC gets closer to the gradient at B.
- Consider the limit as Δx tends to 0.
- This process called differentiation from first principles.
- It gives you **the direction of the steepest uphill (aka. largest increase).**



Gradient/Derived Function, Derivative

- The gradient of the tangent to a curve (non-linear) function $y = f(x)$ varies with variable x. Therefore, it is also a function of x.
- It is called gradient function or derived function.



Gradient/Derived Function, Derivative

- Both $f'(x)$ and $\frac{dy}{dx}$ mean the gradient function.
- Also known as the derivative of y with respect to x.

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

- Practice: how to obtain the general gradient function of $y = x^2$.

$\frac{dy}{dx} = 2x$

Differentiation of Multiple Terms - Polynomials

- A polynomial function: $y = x^3 + 6x^2 - 3x + 1$
- How to differentiate this function with respect to x?

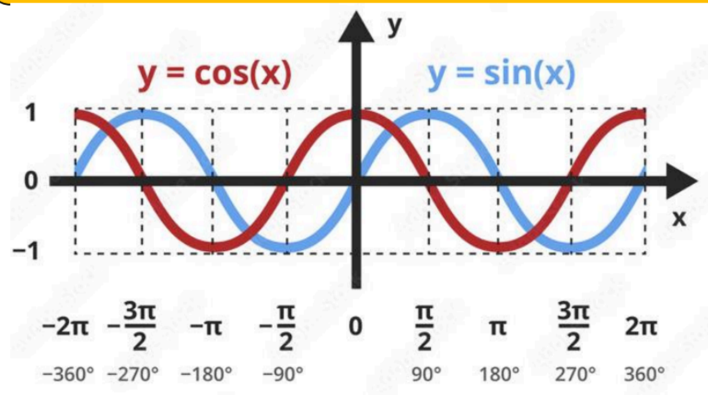
- General rule for sums of functions:

$$\text{If } y = f(x) \pm g(x), \frac{dy}{dx} = f'(x) \pm g'(x)$$

Other Derivatives

- Trigonometric functions: sine and cosine

$$\text{If } f(x) = \sin x, f'(x) = \cos x$$
$$\text{If } f(x) = \cos x, f'(x) = -\sin x$$



Other Derivatives

- Natural exponential

$$\text{If } f(x) = e^x,\ f'(x) = e^x$$

- Natural logarithm (the inverse of the natural exponential)

$$\text{If } f(x) = \ln x \ \ (x > 0),\ f'(x) = \frac{1}{x}$$

The Rules – The Product Rule

$$\text{If y} = f(x)g(x),\ \frac{dy}{dx} = f(x)g'(x) + g(x)f'(x)$$

- Example: $y = x^2 \cos x$

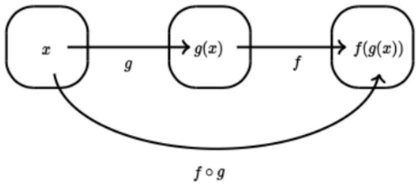
The Rules – The Quotient Rule

$$\text{If y} = \frac{f(x)}{g(x)},\ \frac{dy}{dx} = \frac{g(x)f'(x) - f(x)g\prime(x)}{g(x)^2}$$

- Example: $y = \frac{2x+1}{x^2+2x+1}$

The Rules – The Chain Rule

- Allows us to differentiate a composite function, i.e. a function within a function.
- Composite function:



- How to differentiate it:

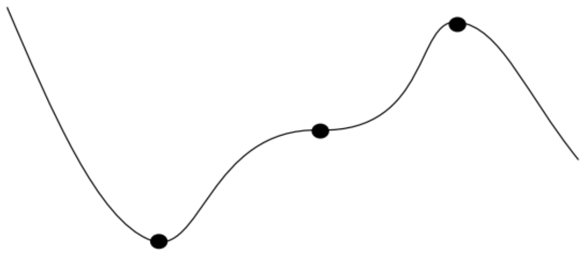
$$\text{If } y = f(g(x))\,,\ \frac{dy}{dx} = f'\big(g(x)\big)g'(x)$$

Outer function differentiated × inner function differentiated

- Example: $y = \ e^{3x}$

How to determine type of stationary point?

- Look at the gradient just before and after the point



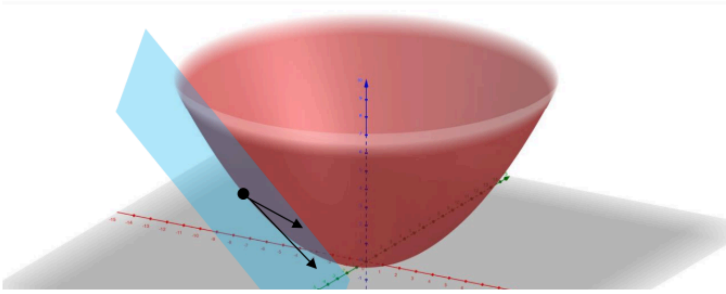
Local Maximum		
Gradient just before	Gradient at max	Gradient just after
?	?	?

Local Minimum		
Gradient just before	Gradient at min	Gradient just after
?	?	?

Point of Inflection		
Gradient just before	Gradient p.o.i	Gradient just after
?	?	?

Multivariate Function and Partial Differentiation

- When a function has more than one independent variable e.g. $z = x^2/10 + y^2/10$, or $f(x, y) = x^2/10 + y^2/10$ 3 dimensions, x and y are independent variables and z is the dependent variable.
- In 3D, a tangent line becomes a tangent plane.
- Directional derivative
- Partial derivative



$z = \frac{x^2}{10} + \frac{y^2}{10}$ or $f(x, y) = \frac{x^2}{10} + \frac{y^2}{10}$

3D space manipulation

Partial Differentiation

Notations: for $z = f(x, y)$, we write:

- $f_x(x, y) = \frac{\partial f}{\partial x} = \frac{\partial z}{\partial x}$, the partial derivative of f with respect to x
- $f_y(x, y) = \frac{\partial f}{\partial y} = \frac{\partial z}{\partial y}$, the partial derivative of f with respect to y

Rule:

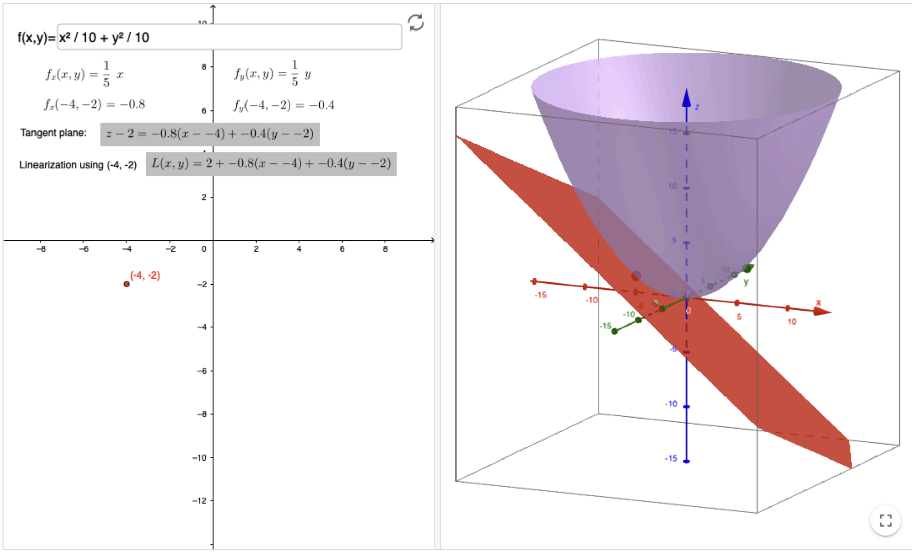
The partial derivative with respect to x is the ordinary derivative of the function of x by treating the other variables as constants.

- To find f_x, treat y as a constant and differentiate $f(x, y)$ with respect to x.
- To find f_y, treat x as a constant and differentiate $f(x, y)$ with respect to y.

UNIVERSITY OF BIRMINGHAM

Example: $f(x, y) = x^2 y + 2x$

( )



Linear functions and the vector notation

- Univariate $\boxed{y = w_0 + w_1 x}$ a line

vector $\boldsymbol{w} = \begin{pmatrix} w_0 \\ w_1 \end{pmatrix}$, $\boldsymbol{x} = \begin{pmatrix} 1 \\ x \end{pmatrix}$

$$\boxed{y = \boldsymbol{w}^T \boldsymbol{x}}$$

- Multivariate $\boxed{y = w_0 + w_1 x_1 + w_2 x_2 .. + w_n x_n}$ a hyperplane

vector $\boldsymbol{w} = \begin{pmatrix} w_0 \\ w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix}$, $\boldsymbol{x} = \begin{pmatrix} 1 \\ x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$

UNIVERSITY OF BIRMINGHAM

$$\boxed{y = \boldsymbol{w}^T \boldsymbol{x}}$$

Revise dot product of a row and a column vectors

First vector transposed to allow proper multiplication

Polynomial function examples

- $y = 3 - 6x + x^2$
- $y = 2 - 3x + 4x^3$
- $y = -1 - 6\sqrt{x} + 4x^2$ (this is not a polynomial function)

Sigmoid (or logistic) function

$$\boxed{y = \frac{1}{1 + e^{-x}}}$$

- Very important function used in machine learning algorithms.
- Homework:
  explore its properties: e.g. shape, output range, symmetry, differentiability.