Calculus for Economics

Chong Chin Herng

May 2023

Contents

1 Limits and Continuity

Limits

Definition 1.1 If the value that a function $f(x)$ approaches as x approaches c is L, then L is known as the **limit** of $f(x)$ as x approaches c and is denoted by

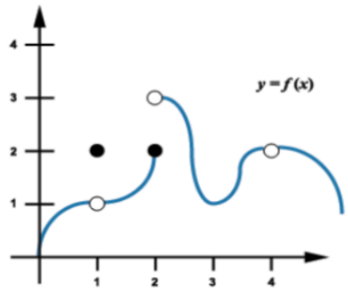
$$\lim_{x \to c} f(x) = L$$

Likewise,

- $\lim_{x \to c^-} f(x)$ is the value that $f(x)$ approaches when x approaches c from the left.
- $\lim_{x \to c^+} f(x)$ is the value that $f(x)$ approaches when x approaches c from the right.

If $\lim_{x \to c^-} f(x) = \lim_{x \to c^+} f(x) = L$, then $\lim_{x \to c} f(x)$ exists and $\lim_{x \to c} f(x) =$ L.

If $\lim_{x \to c^-} f(x) \neq \lim_{x \to c^+} f(x)$, then $\lim_{x \to c} f(x)$ does not exist.



Example 1.2 Figure above shows the graph of $f(x)$. Find $f(1)$, $f(2)$ and $f(4)$. Also, find $\lim_{x \to c} f(x)$ for $c = 1$, $c = 2$ and $c = 4$.

For $c = 1$,

$$\lim_{x \to 1^-} f(x) = \lim_{x \to 1^+} f(x) = 1$$

Therefore,

$$\lim_{x \to 1} f(x) = 1$$

Note that $f(1) = 2$, but $\lim_{x \to 1} f(x) = 1$.

For $c = 2$,

$$\lim_{x \to 2^-} f(x) = 2, \quad \lim_{x \to 2^+} f(x) = 3$$

Since $\lim_{x \to 2^-} f(x) \neq \lim_{x \to 2^+} f(x)$, $\lim_{x \to 2} f(x)$ does not exist. However, $f(2) = 2$.

For $c = 4$,

$$\lim_{x \to 4^-} f(x) = \lim_{x \to 4^+} f(x) = 2$$

Therefore,

$$\lim_{x \to 4} f(x) = 2$$

However, $f(4)$ is undefined.

Theorem 1.3 The following are true:

- If $f(x)$ is a polynomial, then $\lim_{x\to\infty} f(x)$ and $\lim_{x\to-\infty} f(x)$ will tend to either ∞ or $-\infty$.
- If $f(x) = \dfrac{1}{x}$, then $\lim_{x\to\infty} f(x) = \lim_{x\to-\infty} f(x) = 0$. Also, $\lim_{x\to0^-} f(x) = -\infty$ and $\lim_{x\to0^+} f(x) = \infty$, which implies $\lim_{x\to0} f(x)$ does not exist.
- If $f(x) = e^x$, then $\lim_{x\to\infty} f(x) = \infty$ and $\lim_{x\to-\infty} f(x) = 0$.
- If $f(x) = \ln x$, then $\lim_{x\to\infty} f(x) = \infty$ and $\lim_{x\to0^+} f(x) = -\infty$.

Given above are some of the commonly seen limits. Some of these limits might not be trivial to prove, but it is sufficient and extremely useful to visualise these facts by looking at the graphs of these functions using software such as Desmos or GeoGebra.

In fact, the first and second bullet point can be further generalised.

Theorem 1.4 If $f(x) = \dfrac{g(x)}{h(x)}$ where $g(x)$ is a polynomial of degree a and $h(x)$ is a polynomial of degree b, then

- If $a > b$, then $\lim_{x\to\infty} f(x)$ and $\lim_{x\to-\infty} f(x)$ will tend to either ∞ or $-\infty$.
- If $a < b$, then $\lim_{x\to\infty} f(x) = \lim_{x\to-\infty} f(x) = 0$
- If $a = b$, then $\lim_{x\to\infty} f(x) = \lim_{x\to-\infty} f(x) = \dfrac{\alpha}{\beta}$, where α and β are the leading coefficient of $g(x)$ and $h(x)$ respectively.

Notice that the first bullet point in Theorem 1.3 corresponds to the first case in Theorem 1.4, where $a > b = 0$. The second bullet point corresponds to the second case where $0 = a < b$.

We shall illustrate the third case using an example.

Example 1.5 Evaluate $\lim_{x\to\infty} \dfrac{2x^2}{x^2+9}$.

Note that both $2x^2$ and x^2+9 are polynomials of degree 2.

The leading coefficient of a polynomial is the coefficient of the term with the highest power of x in the polynomial. Therefore, the leading coefficient of $2x^2$ is 2 and the leading coefficient of x^2+9 is 1. This gives us the limit

$$\lim_{x\to\infty} \frac{2x^2}{x^2+9} = \frac{2}{1} = 2$$

Continuity

Definition 1.6 A function $f(x)$ is said to be **continuous** at $x = c$ if

$$\lim_{x \to c} f(x) = f(c)$$

Immediately, this implies the following:

- If $\lim_{x \to c} f(x)$ does not exist, then $f(x)$ is not continuous at $x = c$.
- If $f(c)$ is undefined, then $f(x)$ is not continuous at $x = c$.

This means that in Example 1.1, $f(x)$ is not continuous at $x = 1$, $x = 2$ and $x = 4$.

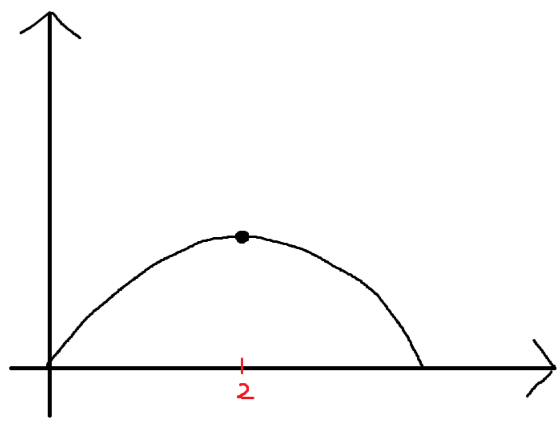


Figure above shows the graph of an example of a function $f(x)$ that is continuous at $x = 2$. Notice that $f(x)$ has the same limit when x approaches 2 from the left and the right. This limit is also the value of $f(2)$. In other words,

$$\lim_{x \to 2^-} f(x) = \lim_{x \to 2^+} f(x) = f(2)$$

Theorem 1.7 All polynomial functions are continuous at every x.

This means that a polynomial function like $f(x) = 15x^2$ or $g(y) = 12y^4$ is continuous everywhere.

Informally, a function is continuous if we can draw the graph of the function without picking up our pen. The function presented in Example 1.1 is not continuous at $x = 1$, $x = 2$ and $x = 4$. Therefore, in an attempt to draw the graph, we have to pick up our pen at $x = 1$, $x = 2$ and $x = 4$. Theorem 1.4 claims that we can draw the graph of any polynomial function in one stroke.

Example 1.8 Given

$$f(x) = \begin{cases} 2x^2 & \text{if } x \leq 1 \\ kx + 6 & \text{if } x > 1 \end{cases}$$

find k such that $f(x)$ is continuous at every x.

Notice that $2x^2$ is a polynomial of degree 2 and $kx + 6$ is a polynomial of degree 1. By Theorem 1.4, both functions are continuous everywhere. The only point at which $f(x)$ might not be continuous is the point where the two polynomials meet, namely $x = 1$.

In other words, we just have to make sure that $f(x)$ is continuous at $x = 1$. We want

$$\lim_{x \to 1^-} f(x) = \lim_{x \to 1^+} f(x) = f(1) = 2(1)^2 = 2$$

In particular, we want

$$\lim_{x \to 1^+} f(x) = 2$$

When considering $\lim_{x \to 1^+} f(x)$, it is important to keep in mind that x is approaching 1 from the right, but is never reaching 1. What we know is that $x > 1$ and x is getting closer and closer to 1. Since $x > 1$, we have

$$\lim_{x \to 1^+} f(x) = \lim_{x \to 1^+} (kx + 6)$$

Since $(kx + 6)$ is a polynomial,

$$\lim_{x \to 1^-} (kx + 6) = \lim_{x \to 1^+} (kx + 6) = \lim_{x \to 1} (kx + 6) = k(1) + 6 = k + 6$$

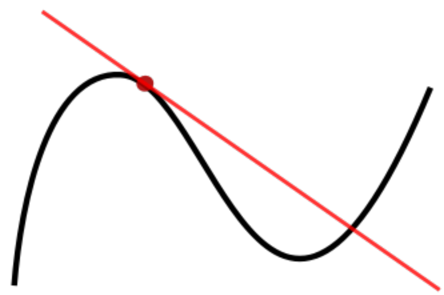
Therefore,

$$\lim_{x \to 1^+} f(x) = k + 6 = 2$$

Solving for k gives us $k = -4$.

2 First Derivative of a Function

Definition 2.1 The **derivative** of a function $f(x)$ at a point $P(a, f(a))$, denoted by $f'(a)$, is the slope[1] of the line that is tangent to the graph of $f(x)$ at P.



This figure from Wikipedia demonstrates the concept of derivative.
The derivative of the graph of this function at the red point is equal to the slope of the red line.

Definition 2.2 The **gradient function** of a function $f(x)$ is a function g where $g(x) = f'(x)$.

Definition 2.3 The process of finding the gradient function of a function is known as **differentiation**.

Notations for the gradient function

There are several ways of denoting a function. Depending on how the function is denoted, its gradient function can also be denoted differently.

- The gradient function of $f(x)$ can be denoted as $f'(x)$.
- The gradient function of a function y can be denoted as y' or $\dfrac{dy}{dx}$.
- A function need not be named. An expression like x^2 can also be treated as a function. We can denote its gradient function by $\dfrac{d}{dx}x^2$.

As a summary, the gradient function of a function $y = f(x)$ is

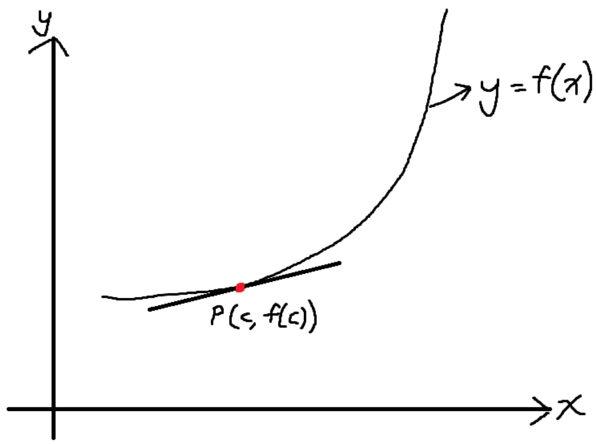
$$y' = f'(x) = \frac{dy}{dx} = \frac{d}{dx}f(x)$$

[1]Slope is also known as gradient. Both terms describe the same concept and only apply to straight lines.

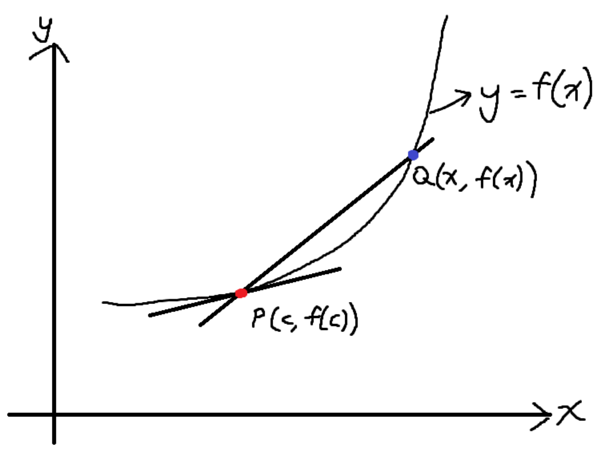
There are other uncommon ways of denoting gradient functions. Notations like these tend to be a good source of confusion for many students. It is important to realise that these are just different ways of denoting the same concept, namely the gradient function.

First Principle of Derivatives

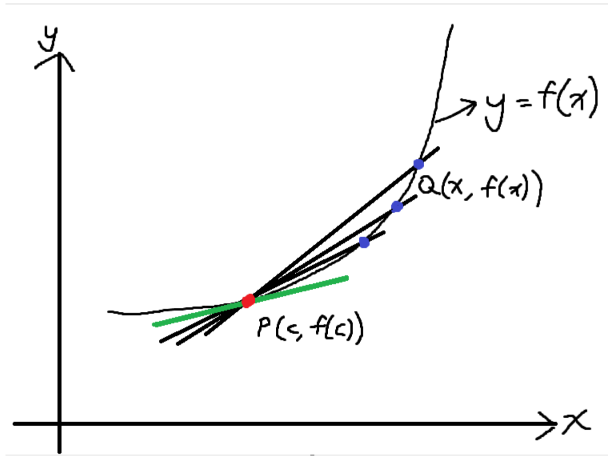
Given a function $f(x)$. We wish to find the slope of the line tangent to the graph of $f(x)$ at the point $P(c, f(c))$.



Instead of the tangent line, consider the line passing through $P(c, f(c))$ and $Q(x, f(x))$.



Imagine moving the point Q closer and closer to P. Note that the line PQ tends to become the tangent line whose slope we wanted to find.



This means that as Q moves closer and closer to P, the slope of the line PQ will be a better and better approximation for the slope of the line tangent to the graph of $f(x)$ at P.

Since $P(c, f(c))$ and $Q(x, f(x))$, the slope of the line PQ will be

$$\frac{f(x) - f(c)}{x - c}$$

As Q moves closer and closer to P, x will tend to c and $f(x)$ will tend to $f(c)$. Therefore, we arrive at the following conclusion

$$f'(c) = \lim_{x \to c} \frac{f(x) - f(c)}{x - c}$$

The right hand side of the above equation is known as the **First Principle of Derivatives**. It is one way of defining derivatives using limits and the very crucial concept behind differential calculus.

Example 2.4 Using the First Principle of Derivatives, differentiate $f(x) = x^2$.

According to the First Principle of Derivatives, the derivative of $f(x)$ at the point $(c, f(c))$ is

$$\begin{aligned} f'(c) &= \lim_{x\to c} \frac{f(x) - f(c)}{x - c} \\ &= \lim_{x\to c} \frac{x^2 - c^2}{x - c} \\ &= \lim_{x\to c} \frac{(x-c)(x+c)}{x - c} \\ &= \lim_{x\to c} (x + c) \\ &= 2c \end{aligned}$$

This means that the gradient function for $f(x)$ is $f'(x) = 2x$.

Differentiation Formulae and Rules

Using the First Principle of Derivative, mathematicians were able to prove formulae and rules that simplifies the process of differentiation. The theorem below summarises some of the useful ones.

Theorem 2.5 The following are true:

- $\dfrac{d}{dx}k = 0 \quad$ for any constant k.
- $\dfrac{d}{dx}x^n = nx^{n-1} \quad$ for any constant n.
- $\dfrac{d}{dx}e^x = e^x$
- $\dfrac{d}{dx}\ln x = \dfrac{1}{x}$

Example 2.6 Differentiate $f(x) = 9$ and $g(x) = x^9$.

$$\text{Since 9 is a constant, we have } f'(x) = 0 \text{ and } g'(x) = 9x^{9-1} = 9x^8$$

Example 2.7 Find the slope of the line that is tangent to the graph of $f(x) = 9$ at the points $(1, 9)$ and $(100, 9)$.

$$\text{The question essentially asks for the value of } f'(1) \text{ and } f'(100)$$

$$\text{From Example 2.2, note that } f'(x) = 0 \text{ regardless of } x$$

$$\text{Therefore, } f'(1) = f'(100) = 0$$

Example 2.8 Find the slope of the line that is tangent to the graph of $g(x) = x^9$ at the points $(1, 1)$ and $(2, 512)$.

$$\text{The question essentially asks for the value of } g'(1) \text{ and } g'(2)$$

$$\text{From Example 2.2, } g'(x) = 9x^8$$

$$\text{Therefore, } g'(1) = 9(1)^8 = 9 \text{ and } g'(2) = 9(2)^8 = 2304$$

Theorem 2.9 The following are true:

- If $f(x) = g(x) + h(x)$, then $f'(x) = g'(x) + h'(x)$
- If $f(x) = g(x) - h(x)$, then $f'(x) = g'(x) - h'(x)$
- If $f(x) = kg(x)$, then $f'(x) = kg'(x)$
- **Product Rule:** If $f(x) = g(x)h(x)$, then $f'(x) = g'(x)h(x) + g(x)h'(x)$
- **Quotient Rule:** If $f(x) = \frac{g(x)}{h(x)}$, then $f'(x) = \frac{g'(x)h(x) - g(x)h'(x)}{[h(x)]^2}$
- **Chain Rule:** If $f(x) = g(h(x))$, then $f'(x) = g'(h(x))h'(x)$

Example 2.10 Differentiate $f(x) = x^2 + 1$.

$$\begin{aligned} f'(x) &= \frac{d}{dx}x^2 + \frac{d}{dx}1 \\ &= 2x + 0 \\ &= 2x \end{aligned}$$

Example 2.11 Differentiate $f(x) = e^x - x$.

$$\begin{aligned} f'(x) &= \frac{d}{dx}e^x - \frac{d}{dx}x \\ &= e^x - 1 \end{aligned}$$

Example 2.12 Differentiate $f(x) = 5 \ln x$.

$$\begin{aligned} f'(x) &= 5\frac{d}{dx}\ln x \\ &= 5(\frac{1}{x}) \\ &= \frac{5}{x} \end{aligned}$$

Example 2.13 Differentiate $f(x) = x^5 e^x$.

$$\begin{aligned} f'(x) &= e^x \frac{d}{dx}x^5 + x^5 \frac{d}{dx}e^x \\ &= 5x^4 e^x + x^5 e^x \end{aligned}$$

Example 2.14 Differentiate $f(x) = \dfrac{\ln x}{e^x}$.

$$\begin{aligned} f'(x) &= \frac{e^x \frac{d}{dx} \ln x - \ln x \frac{d}{dx} e^x}{(e^x)^2} \\ &= \frac{\frac{e^x}{x} - e^x \ln x}{e^{2x}} \\ &= \frac{e^x - xe^x \ln x}{xe^{2x}} \\ &= \frac{e^x(1 - x\ln x)}{xe^{2x}} \\ &= \frac{1 - x\ln x}{xe^x} \end{aligned}$$

Example 2.15 Differentiate $f(x) = (\ln x)^3$.

Let $g(x) = x^3$ and $h(x) = \ln x$. Note that $f(x) = g(h(x)) = (\ln x)^3$

We have $g'(x) = 3x^2$, therefore $g'(h(x)) = 3(\ln x)^2$. Moreover, $h'(x) = \dfrac{1}{x}$

Therefore, $f'(x) = g'(h(x))h'(x) = \dfrac{3(\ln x)^2}{x}$

Implicit Differentiation

Instead of differentiating a function y in terms of x to obtain $\dfrac{dy}{dx}$, we might sometimes be given an equation relating x and y like $x^2 + y^2 = 25$. To find $\dfrac{dy}{dx}$ for such equations, it is often impractical or impossible to solve for y before differentiating the resulting function. A simpler way to compute $\dfrac{dy}{dx}$ in this case will be to perform an implicit differentiation.

The general process of implicit differentiation consists of two steps:

1. Differentiate both sides of the given equation with respect to x.
2. Solve for $\dfrac{dy}{dx}$.

Example 2.16 Given $x^2 + y^2 = 25$, find $\dfrac{dy}{dx}$.

For this equation, it is actually possible to solve for y.

$$y = \pm\sqrt{25 - x^2}$$

We can then differentiate y by splitting into two cases. When $y = \sqrt{25 - x^2}$,

$$\frac{dy}{dx} = \frac{d}{dx}(25 - x^2)^{\frac{1}{2}} = \frac{1}{2}(25 - x^2)^{-\frac{1}{2}}(-2x) = -\frac{x}{(25 - x^2)^{\frac{1}{2}}} = -\frac{x}{\sqrt{25 - x^2}}$$

Likewise, when $y = -\sqrt{25 - x^2}$, we have

$$\frac{dy}{dx} = \frac{d}{dx}(-\sqrt{25 - x^2}) = -\frac{d}{dx}(25 - x^2)^{\frac{1}{2}} = -\left(-\frac{x}{\sqrt{25 - x^2}}\right) = \frac{x}{\sqrt{25 - x^2}}$$

Now we shall perform implicit differentiation to obtain the same result. First, we differentiate both sides of the equation $x^2 + y^2 = 25$ with respect to x.

$$\frac{d}{dx}(x^2 + y^2) = \frac{d}{dx}25$$

$$2x + 2y\,\frac{dy}{dx} = 0$$

Note that since y is a function of x, we used chain rule to differentiate y^2 with respect to x. We first differentiate y^2 with respect to y, then multiply the result with $\dfrac{dy}{dx}$.

We can now solve for $\dfrac{dy}{dx}$, obtaining

$$\frac{dy}{dx} = -\frac{x}{y}$$

Indeed, when $y = \sqrt{25 - x^2}$,

$$-\frac{x}{y} = -\frac{x}{\sqrt{25 - x^2}}$$

When $y = -\sqrt{25 - x^2}$,

$$-\frac{x}{y} = -\frac{x}{-\sqrt{25 - x^2}} = \frac{x}{\sqrt{25 - x^2}}$$

This result agrees with what we obtained in the first method.

Example 2.17 Given $3x^2 - 5y^2 + 9x = 25 - 15y$, find $\dfrac{dy}{dx}$.

Differentiating both sides with respect to x,

$$6x - 10y\,\frac{dy}{dx} + 9 = 0 - 15\frac{dy}{dx}$$

Now we solve for $\dfrac{dy}{dx}$, obtaining

$$\begin{aligned} 6x + 9 &= 10y\,\frac{dy}{dx} - 15\frac{dy}{dx} \\ \frac{dy}{dx}(10y - 15) &= 6x + 9 \\ \frac{dy}{dx} &= \frac{6x + 9}{10y - 15} \end{aligned}$$

We need not substitute y into the result, since we could not even solve for y to begin with.

Interpretations

The slope of a straight line describes the rate of change of the straight line. Derivatives generalise the notion of rate of change to an arbitrary function. Unlike straight lines, the rate of change of a curve is not a constant. Derivatives make use of the slope of the tangent lines to different points on the curve to describe its rate of change.

It is important to keep in mind that the derivative of a function $f(x)$ at any given point $(a, f(a))$ is a number $f'(a)$. This number is the slope of the line tangent to the function $f(x)$ at the point $(a, f(a))$.

When we are differentiating a function $f(x)$, we obtain the gradient function $f'(x)$. This gradient function is also sometimes known as the derivative of $f(x)$. In this case, the derivative is no longer defined as a number, but as a function. This gradient function is a function that, when evaluated at $x = a$, gives us the derivative of $f(x)$ at the point $(a, f(a))$. This concept is demonstrated in Example 2.7 and 2.8.

When we say that the derivative of $f(x)$ at $x = a$ is $f'(a)$, this means that at the point $(a, f(a))$, when x changes by a small amount dx, the $f(x)$ should change by $f'(a)dx$. For example, if $f'(3) = 2$, then at the point $(3, f(3))$, when x changes by 0.001 (so that it is now 3.001), $f(x)$ should change by approximately[2] 0.002 (so that it is now $f(3) + 0.002$).

Definition 2.18 The **differential** of a function $f(x)$, denoted by df, is the change in the value of $f(x)$ when x changes by a small amount dx.

Theorem 2.19 Let $f(x)$ be a function, then

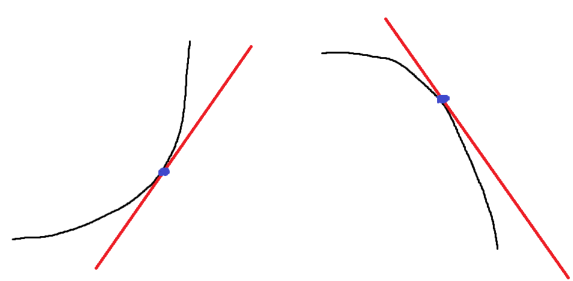
$$df = f'(x)dx$$

Example 2.20 Find the differential of $f(x) = x^3$. Hence, approximate $(2.001)^3$.

Since $f'(x) = 3x^2$, we have $df = 3x^2dx$. To approximate $(2.001)^3$, we pick the point where $x = 2$ and analyse how $f(x)$ changes when x changes by 0.001. Since $f'(2) = 3(2)^2 = 12$, we know that at the point $(2, f(2))$, $f(x)$ should change by approximately 0.012 when x changes by 0.001. Note that $f(2) = 2^3 = 8$, therefore we approximate $(2.001)^3$ to be $8 + 0.012 = 8.012$.

Indeed, $(2.001)^3 = 8.012006$.

We can use the derivative of the function at any given point $(a, f(a))$ to determine whether the function is increasing or decreasing at that point. Figure below demonstrates this concept.



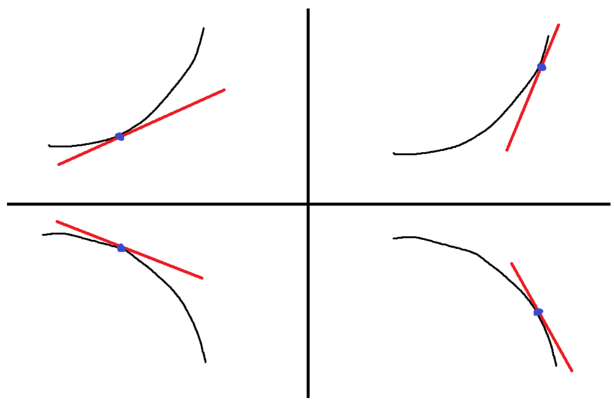
As shown on the left, if the function is increasing at the blue point, the line tangent to the function at that point will rise from left to right, thus will have

[2] We mentioned that x should only change by a small amount, but we did not mention how small. In fact, x should change by an infinitely small amount. 0.001 is far from being infinitely small, therefore the change in $f(x)$ cannot be exactly, but is very close to 0.002.

a positive slope. Similarly, as shown on the right, if the function is decreasing at the blue point, the derivative of the function at that point will be negative.

Theorem 2.21 If $f'(a)$ is positive, then $f(x)$ is increasing at the point $(a, f(a))$. If $f'(a)$ is negative, then $f(x)$ is decreasing at the point $(a, f(a))$.

Therefore, the sign of the derivative tells us whether a function is increasing or decreasing at a point. On the other hand, the magnitude of the derivative shows the sensitivity of change of a function's output with respect to the input. Figure below demonstrates this concept.



The top-left and the top-right figure both show a positive derivative. However, the magnitude of the derivative of the top-right figure will be larger than that of the top-left figure, and hence we see that the tangent line in the top-right figure is much steeper than that of the top-left figure.

Similarly, both derivatives in the bottom-left and bottom-right figures are negative. However, the magnitude of the derivative of the bottom-right figure tend to be larger (more negative), and hence the tangent line appear to be steeper as compared to the bottom-left figure.

When the graph of a function increases then decreases, a maximum point is formed at the point where the graph changes from being increasing to decreasing. Similarly, when the graph of a function decreases then increases, a minimum point is formed at the point where the graph changes from being decreasing to increasing.

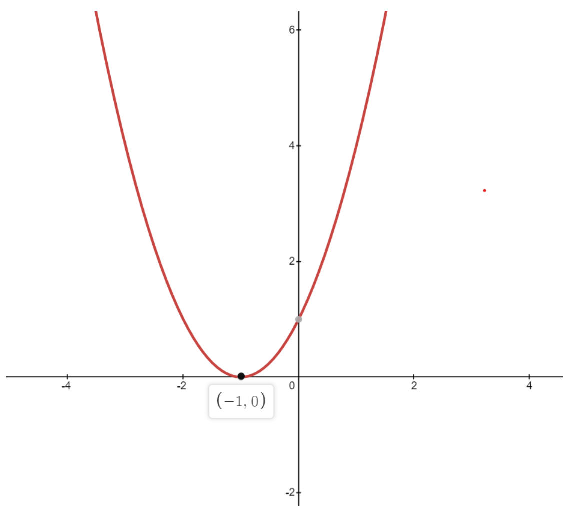
By Theorem 2.15, when the graph of the function increases then decreases, its derivative will be positive then negative. When the graph of the function decreases then increases, its derivative will be negative then positive. At minimum or maximum points of a graph, it makes sense that the graph should have a derivative of zero. Indeed, when $f'(c) = 0$, it is very likely that $f(x)$ has a maximum or minimum point at $x = c$.

Example 2.22 Find the minimum point of the function $f(x) = x^2 + 2x + 1$.

Differentiating $f(x)$ gives $f'(x) = 2x + 2$. When $f'(x) = 0$, $x = -1$. This means that $f(x)$ very likely has a maximum or minimum point at $x = -1$. Notice that when $x < -1$, $f'(x) < 0$. Therefore, the $f(x)$ is decreasing for $x < -1$. On the other hand, when $x > -1$, $f'(x) > 0$, so $f(x)$ is increasing for $x > -1$.

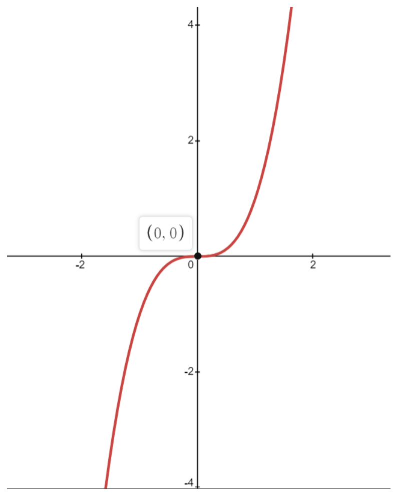
Now we know that $f(x)$ is decreasing for $x < -1$ and increasing for $x > -1$. This means that $f(x)$ attains a minimum point at $x = -1$. Evaluating $f(x)$ at -1 gives $f(-1) = (-1)^2 + 2(-1) + 1 = 0$. This means that $f(x)$ has a minimum point at $(-1, 0)$.

Indeed, this is the graph of $f(x)$ generated using Desmos.



Example 2.23 Find the maximum or minimum points of the function $f(x) = x^3$.

Note that $f'(x) = 3x^2$. When $f'(x) = 0$, $x = 0$. Hence $f(x)$ very likely has a minimum or maximum point at $x = 0$. However, note that $x < 0$ or $x > 0$, $f'(x) = 3x^2 > 0$, since the square of any nonzero number is positive. This means that $f(x)$ increases for $x < 0$, stops increasing at $x = 0$, then increases again for $x > 0$. Neither does $f(x)$ attain a maximum nor a minimum point at $x = 0$. Figure below is the graph of $f(x)$.



Example 2.24 Find x that gives the maximum and minimum value of $f(x) = 2x^3 - 21x^2 + 60x - 24$ where $1 \leq x \leq 7$.

$f'(x) = 6x^2 - 42x + 60$. When $f'(x) = 0$,

$$6x^2 - 42x + 60 = 0$$
$$6(x^2 - 7x + 10) = 0$$
$$(x-2)(x-5) = 0$$

Now we know that when $x = 2$ or $x = 5$, $f'(x) = 0$. Note that

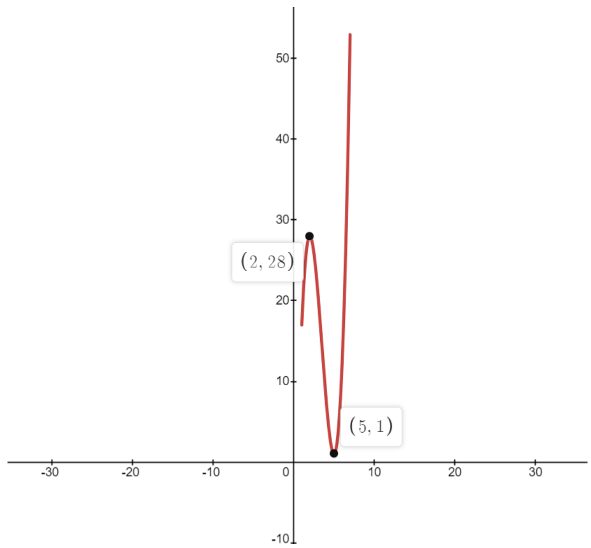
- When $x < 2$, both $(x-2)$ and $(x-5)$ will be negative, thus $f'(x) = (x-2)(x-5)$ will be positive.
- When $2 < x < 5$, $(x-2)$ is positive but $(x-5)$ is negative, thus $f'(x) = (x-2)(x-5)$ is negative.
- When $x > 5$, both $(x-2)$ and $(x-5)$ will be positive, thus $f'(x) = (x-2)(x-5)$ will be positive.

Additionally, note that $f(1) = 17$, $f(2) = 28$, $f(5) = 1$ and $f(7) = 53$.

This means that from $f(x)$ starts from the point $(1, 17)$, increases to $(2, 28)$, then decreases to $(5, 1)$, then increases back to $(7, 53)$.

Therefore, for $1 \leq x \leq 7$, the maximum value of $f(x)$ has to be 53, and the minimum value of $f(x)$ has to be 1.

It is not mandatory but always satisfying to check the graph of $f(x)$ to make sure maths is working.



Profit Maximising Condition

The **cost function**, commonly denoted as $C(q)$, measures the minimum cost for an organisation to produce a given quantity q of output. The **marginal cost** $MC = C'(q)$ is the derivative of the cost function. Marginal cost measures the change in cost that comes from producing one additional unit of output.

Besides cost function, an organisation can also analyse its **revenue function**. The revenue function $R(q)$ measures the revenue from selling a given quantity q of output. Likewise, the **marginal revenue** $MR = R'(q)$ is the derivative of the revenue function, and it measures how much in revenue an organisation earns for each additional unit of output sold.

In general, differentiation can be used to find marginals. Analysis of marginals allows organisations to optimise production and overall operations. One classic example of this is the **profit maximising condition** of $MR = MC$. The profit for an organisation is maximised when the value of the last unit of product (marginal revenue) equals the cost of producing the last unit of production (marginal cost).

Typically, the marginal cost of producing a small amount of output will be lower than the corresponding marginal revenue. We continue to produce output when $MR > MC$ so that we earn the resulting profit. Eventually, we reach the point where $MR = MC$, after which MC will grow to be larger than MR, causing us to lose profit should we continue to produce output. The rule of thumb here is to produce if $MR > MC$.

Price Elasticity of Demand

Economists employ price elasticity of demand to understand how supply and demand change when a product's price changes. Here, supply and demand are often denoted as Q, and price is often denoted as P.

Given a product, we can analyse its price elasticity of demand from its demand function $Q(P)$. The **price elasticity of demand** of a product at a certain point of production $(P, Q(P))$ is defined to be the ratio of the percentage change in Q to the percentage change in P. Denote dQ to be change in Q and dP to be change in P, then the formula for price elasticity of demand at $(P, Q(P))$, denoted as e_P, can be written as

$$e_P = \frac{\frac{dQ}{Q} \cdot 100\%}{\frac{dP}{P} \cdot 100\%}$$

The formula can be further simplified and rewritten as

$$e_P = \frac{dQ}{dP} \cdot \frac{P}{Q} = \frac{P}{Q} Q'(P)$$

This means that given the demand function for a product, we can find its price elasticity of demand at a certain point of production by differentiating its demand function.

In general, the term "elasticity" in economics is used to describe the ratio of the percentage change in a function's value to the percentage change in its variables. Given a function $y = f(x)$, the elasticity of y with respect to x is

$$e_y = \frac{dy}{dx} \cdot \frac{x}{y} = \frac{x}{y} f'(x)$$

Example 2.21 Given the demand function for a product $Q(P) = 800 - 50P^2$, find the price elasticity of demand when 50 units of the product are produced.

We differentiate the demand function, obtaining $Q'(P) = -100P$. Since 50 units of the product are produced, our point of production should have $Q(P) = 50$. We can solve for P as follows

$$\begin{aligned} 50 &= 800 - 50P^2 \\ 50P^2 &= 750 \\ P^2 &= 15 \\ P &= \sqrt{15} \end{aligned}$$

Therefore, the price of elasticity at this point of production would be $\frac{P}{Q} Q'(P) = \frac{\sqrt{15}}{50}(-100\sqrt{15}) = -30$.

Generally, when interpreting price of elasticity, we ignore its sign and only consider its magnitude. When the magnitude of price of elasticity at a point of production is greater than 1, we say that the demand is elastic at that point of production. Otherwise, the demand is said to be inelastic at that point of production. Therefore, in Example 2.21, since the price of elasticity has a magnitude of 30, the product is considered to be elastic at the point of production $(\sqrt{15}, 50)$.

3 Second Derivative of a Function

Definition 3.1 The **second derivative** of a function $f(x)$ at a point $(a, f(a))$, denoted by $f''(a)$, is the derivative of $f'(x)$ at the point $(a, f'(a))$.

In the above definition, the second derivative is defined as a number representing the slope of the line tangent to the gradient function of $f(x)$ at a given point. We tend to also define the second derivative of $f(x)$ as the gradient function of $f'(x)$. In this case, the second derivative is defined to be a function rather than a number. Informally, the second derivative can be phrased as "the rate of change of the rate of change".

Notations for the gradient function

Similar to the (first) derivative of a function, there are also several notations used to denote the second derivative of a function. In general, for a function $y = f(x)$, its second derivative as be written as the following.

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

Example 3.2 Find the second derivative of $f(x) = -\dfrac{5}{x}$.

Note that $f(x) = -5x^{-1}$. Therefore,

$$\begin{aligned} f'(x) &= 5x^{-2} \\ f''(x) &= -10x^{-3} \\ &= -\frac{10}{x^3} \end{aligned}$$

Example 3.3 Find the second derivative of $y = 2e^{4x}$.

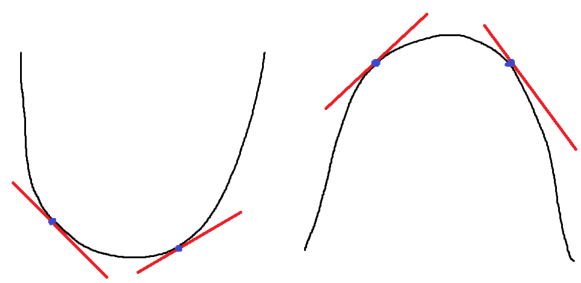
$$\begin{aligned} \frac{dy}{dx} &= 2(4)e^{4x} \\ &= 8e^{4x} \\ \frac{d^2y}{dx^2} &= 8(4)e^{4x} \\ &= 32e^{4x} \end{aligned}$$

Example 3.4 Find $\dfrac{d^2}{dx^2}e^{7x-4}$.

$$\begin{aligned}\frac{d}{dx}e^{7x-4} &= 7e^{7x-4}\\ \frac{d^2}{dx^2}e^{7x-4} &= 7(7)e^{7x-4}\\ &= 49e^{7x-4}\end{aligned}$$

Interpretations

The second derivative of a function measures the rate of change of the slope of the line tangent to the curve of the function. It turns out that we can use the second derivative to determine whether a curve is concave-up or concave-down at a given point. Figure below illustrates concavity of a curve.



We say that a curve is concave-up at a point P if the line tangent to the curve at P is below the curve. Similarly, a curve is concave-down at a point P if the line tangent to the curve at P is above the curve.

There are important observations to be made on the figure above.

- When the curve is concave-up, its first derivative (the slope of the red line) increases from being negative to positive as we move from left to right.
- Similarly, when the curve is concave-down, its first derivative decreases from being positive to negative.

In the previous section, we learnt that if a function is increasing, its first derivative will be positive. If a function is decreasing, its first derivative will be negative. What happens if the first derivative itself is increasing?

Since the second derivative is the first derivative of the first derivative, if the first derivative is increasing, the second derivative will be positive! Likewise, if the first derivative of the function is decreasing, the second derivative of the function will be negative.

From the figure above, we showed that if the first derivative of a function at a point is increasing, then the function is concave-up at that point, and vice-versa. We now arrive at the following important result.

Theorem 3.5 If $f''(a)$ is positive, then $f(x)$ is concave-up at point $(a, f(a))$. If $f''(a)$ is negative, then $f(x)$ is concave-down at point $(a, f(a))$.

4 Indefinite Integration

Definition 4.1 The **antiderivative**, also known as the **indefinite integral** of a function $f(x)$ with respect to x, denoted as

$$\int f(x)\, dx$$

is a function $F(x)$ such that $F'(x) = f(x)$.

Definition 4.2 The process of finding the antiderivatives of a function is known as **indefinite integration**.

Indefinite Integration Formulae and Rules

Theorem 4.3 The following are true:

- $\displaystyle\int 0\, dx = C \quad$ for any constant C.
- $\displaystyle\int x^n\, dx = \frac{x^{n+1}}{n+1} + C \quad$ for any constant $n \neq -1$ and C.
- $\displaystyle\int e^{ax}\, dx = \frac{1}{a} e^{ax} + C \quad$ for any constant $a \neq 0$ and C.
- $\displaystyle\int \frac{1}{x}\, dx = \ln|x| + C \quad$ for any constant C.

Example 4.4 Find $\displaystyle\int x\sqrt{x}\, dx$.

Note that $\sqrt{x} = x^{\frac{1}{2}}$. Therefore,

$$\begin{aligned} \int x\sqrt{x}\, dx &= \int x^{\frac{3}{2}}\, dx \\ &= \frac{x^{\frac{5}{2}}}{\frac{5}{2}} + C \\ &= \frac{2x^{\frac{5}{2}}}{5} + C \end{aligned}$$

where C is any constant.

Example 4.5 Find $\displaystyle\int \frac{1}{\sqrt[5]{x}}\,dx$.

Note that $\dfrac{1}{\sqrt[5]{x}} = \dfrac{1}{x^{\frac{1}{5}}} = x^{-\frac{1}{5}}$. Therefore,

$$\begin{aligned}\int \frac{1}{\sqrt[5]{x}}\,dx &= \int x^{-\frac{1}{5}}\,dx \\ &= \frac{x^{\frac{4}{5}}}{\frac{4}{5}} + C \\ &= \frac{5x^{\frac{4}{5}}}{4} + C\end{aligned}$$

where C is any constant.

Theorem 4.6 The following are true:

- $\displaystyle\int f(x) + g(x)\,dx = \int f(x)\,dx + \int g(x)\,dx$
- $\displaystyle\int f(x) - g(x)\,dx = \int f(x)\,dx - \int g(x)\,dx$
- $\displaystyle\int kf(x)\,dx = k\int f(x)\,dx$

Example 4.7 Find $\displaystyle\int 6x^5 - 18x^2 + 7\,dx$.

$$\begin{aligned}\int 6x^5 - 18x^2 + 7\,dx &= \int 6x^5\,dx - \int 18x^2\,dx + \int 7\,dx \\ &= 6\int x^5\,dx - 18\int x^2\,dx + 7\int x^0\,dx \\ &= 6\left(\frac{x^6}{6}\right) - 18\left(\frac{x^3}{3}\right) + 7\left(\frac{x^1}{1}\right) + C \\ &= x^6 - 6x^3 + 7x + C\end{aligned}$$

where C is any constant.

Example 4.8 Find $\int 10w^4 + 9w^3 + 7w\,dw$.

$$\begin{aligned}\int 10w^4 + 9w^3 + 7w\,dw &= 10\int w^4\,dw + 9\int w^3\,dw + 7\int w\,dw \\ &= 10\left(\frac{w^5}{5}\right) + 9\left(\frac{w^4}{4}\right) + 7\left(\frac{w^2}{2}\right) + C \\ &= 2w^5 + \frac{9}{4}w^4 + \frac{7}{2}w^2 + C\end{aligned}$$

where C is any constant.

Integration by Substitution

One important fact about calculus is that it is generally more difficult to evaluate integrals of functions than to differentiate functions. In other words, differentiation is generally easier than integration.

This stems from the fact that in differentiation, it is possible to differentiate any function via repeated applications of product rule, quotient rule and chain rule. However, no equivalent forms for these three rules exist in integration. Therefore, not all functions are trivial to integrate, and several techniques have been developed over time to help us evaluate more complicated integrals.

Integration by substitution, also known as **u-Substitution** or **The Reverse Chain Rule**, is one such technique. This technique corresponds to the chain rule of differentiation, but unlike the chain rule, there is no formula we can write down that allows us to apply it in a straightforward manner.

Here is a demonstration of how integration by substitution works in the general case. Whenever we are evaluating an integral of the form

$$\int f'(g(x))g'(x)\,dx$$

We let $u = g(x)$, therefore $du = g'(x)dx$. Our integral simplifies to

$$\int f'(u)\,du = f(u) + C = f(g(x)) + C$$

where C is a constant. Indeed, this result is consistent with the chain rule of differentiation:

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

Now we consider some actual examples.

Example 4.9 Find $\int 2xe^{x^2}\,dx$.

Let $f(x) = e^x$ and $g(x) = x^2$. Then, $f'(x) = e^x$ and $g'(x) = 2x$. Notice that our integral is of the form

$$\int f'(g(x))g'(x)\,dx$$

We let $u = x^2$, therefore $du = 2xdx$. Our integral simplifies to

$$\int e^u\,du = e^u + C = e^{x^2} + C$$

where C is a constant.

The difficult part of performing integration by substitution is to recognise that an integral is in the correct form. It takes practice and experience to build intuitions into recognising such patterns. In general, we are looking for a function $g(x)$ in the integrand[3], such that $g'(x)dx$ is also in the integrand. In the above example, we observed that x^2 and $2x\,dx$ are present in the integrand, therefore we substituted $u = x^2$ to simplify the integral.

Sometimes, the integral might not be given in the correct form, but can be transformed into the correct form via some manipulations.

Example 4.10 Find $\int \frac{x}{x^2+1}\,dx$.

The derivative of $x^2 + 1$ is $2x$. However, there is no $2x\,dx$ in the integrand. We can fix this by multiplying and dividing the integrand by 2.

$$\int \frac{x}{x^2+1}\,dx = \frac{1}{2}\int \frac{2x}{x^2+1}\,dx = \frac{1}{2}\int \frac{1}{x^2+1}2x\,dx$$

Now, we realise that $x^2 + 1$ and $2x\,dx$ are in the integrand. Therefore, we substitute $u = x^2 + 1$, so $du = 2x\,dx$.

$$\frac{1}{2}\int \frac{1}{x^2+1}2x\,dx = \frac{1}{2}\int \frac{1}{u}du = \frac{1}{2}\ln u + C = \frac{1}{2}\ln(x^2+1) + C$$

where C is a constant.

[3]The integrand is the part of the integral between the stylish S and the dx. It is the function that we are integrating.

Sometimes, we need not use integration by substitution to evaluate an integral, but using the technique may simplify the evaluation.

Example 4.11 Find $\int (x+1)^3 \, dx$.

One method of evaluating this integral is to expand $(x+1)^3$, giving us

$$\int (x+1)^3 \, dx = \int x^3 + 3x^2 + 3x + 1 \, dx$$

We can then integrate each term separately without using any substitutions.

However, we can also notice that the derivative of $x+1$ is 1, and

$$\int (x+1)^3 \, dx = \int (x+1)^3 \cdot 1 \, dx$$

Since $x+1$ and $1 \, dx$ are both present in the integrand, we substitute $u = x+1$, so $du = dx$, giving us

$$\int (x+1)^3 \cdot 1 \, dx = \int u^3 du = \frac{u^4}{4} + C = \frac{1}{4}(x+1)^4 + C$$

where C is a constant.

Integration by Parts

Integration by parts is an integration technique that corresponds to the product rule of differentiation.

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

We take the indefinite integral on both sides, giving us

$$\int \frac{d}{dx} f(x)g(x) \, dx = \int f'(x)g(x) \, dx + \int f(x)g'(x) \, dx$$

Recall that the indefinite integral of a function is the antiderivative of the function. On the left hand side of the equation, we are taking the derivative of $f(x)g(x)$, then taking its antiderivative. This gives us the original function $f(x)g(x)$.

$$f(x)g(x) = \int f'(x)g(x) \, dx + \int f(x)g'(x) \, dx$$

Let $u = f(x)$ and $v = g(x)$. Then, $du = f'(x)dx$ and $dv = g'(x)dx$. Substituting these equations give us

$$uv = \int v \, du + \int u \, dv$$

Rearranging the equation gives us the classic formula for integration by parts.

$$\int u\,dv = uv - \int v\,du$$

Example 4.12 Find $\displaystyle\int xe^x\,dx$.

The usual first step in applying integration by parts is to split the integrand into u and dv. In other words, we need to split the integrand into $f(x)$ and $g'(x)dx$. In this case, we set $f(x) = x$ and $g'(x) = e^x$. This means $u = x$ and $dv = e^x dx$. Notice that our integral is now in the form

$$\int u\,dv$$

We can then deduce du and v from u and dv. Since $u = x$, we have $du = dx$. From $dv = e^x dx$, we integrate both sides to get $v = e^x$. We can now directly apply the integration by parts formula.

$$\int xe^x\,dx = \int u\,dv = uv - \int v\,du = xe^x - \int e^x dx = xe^x - e^x + C$$

where C is a constant.

Similar to integration by substitution, it takes practice and experience to be able to identify whether an integral can be evaluated with the help of integration by parts. The choice of u and v will often not be obvious.

Example 4.13 Find $\displaystyle\int \ln x\,dx$.

By viewing $\ln x$ as $1 \cdot \ln x$, we can set $u = \ln x$ and $dv = 1\,dx$. This implies $du = \dfrac{1}{x}dx$ and $v = x$.

$$\int \ln x\,dx = x\ln x - \int x\frac{1}{x}\,dx = x\ln x - \int 1\,dx = x\ln x - x + C$$

where C is a constant.

Example 4.14 Find $\displaystyle\int x \ln x \, dx$.

Set $u = \ln x$ and $dv = x\,dx$. Therefore, $du = \dfrac{1}{x}dx$ and $v = \dfrac{x^2}{2}$.

$$\int x \ln x \, dx = \frac{x^2}{2}\ln x - \int \frac{x^2}{2}\frac{1}{x}\,dx = \frac{1}{2}x^2 \ln x - \int \frac{x}{2}\,dx = \frac{1}{2}x^2\ln x - \frac{1}{4}x^2 + C$$

where C is a constant.

5 Definite Integration

In calculus, we use differentiation to find the slope of the tangent line to a curve at different points. On the other hand, integration can help us to find the area under a curve.

The **Fundamental Theorem of Calculus** connects the notion of antiderivatives of a function with the area under the curve of the function. The theorem explains why evaluating antiderivatives of functions helps us to find the area under the function's curve.

Theorem 4.15 (Fundamental Theorem of Calculus) [4] Let $F(x)$ be the antiderivative of a function $f(x)$. In other words, $\int f(x)\,dx = F(x)$. Let A be the area under the curve of $f(x)$ from $x = a$ to $x = b$. Then,

$$A = F(b) - F(a)$$

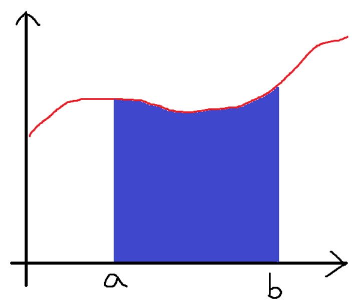


Figure above demonstrates what is meant by "area under the curve from $x = a$ to $x = b$". We normally assume that a and b are real numbers where $a \leq b$, and that $f(x)$ is non-negative when x is in the range $[a, b]$.

Definition 4.16 The area under the curve of $f(x)$ from $x = a$ to $x = b$ is known as the **definite integral** of $f(x)$ from a to b, and is denoted as

$$\int_a^b f(x)\,dx$$

This means that definite integrals are numbers, while indefinite integrals or antiderivatives are functions.

[4] The full theorem is split into two parts. This is a non-rigorous description of the full theorem. Several subtle details regarding differentiability and continuity are omitted.

Example 4.17 Find $\int_0^4 x\,dx$.

The antiderivative of $f(x) = x$ is $F(x) = \frac{1}{2}x^2$. Applying the Fundamental Theorem of Calculus, we have

$$\int_0^4 x\,dx = \frac{1}{2}4^2 - \frac{1}{2}0^2 = 8$$

In the context of evaluating definite integrals, we usually denote $F(b) - F(a)$ as $[F(x)]_a^b$. Therefore, we can also write

$$\int_0^4 x\,dx = \left[\frac{1}{2}x^2\right]_0^4 = 8$$

This means that the area under the curve[5] $y = x$ from $x = 0$ to $x = 4$ is 8.

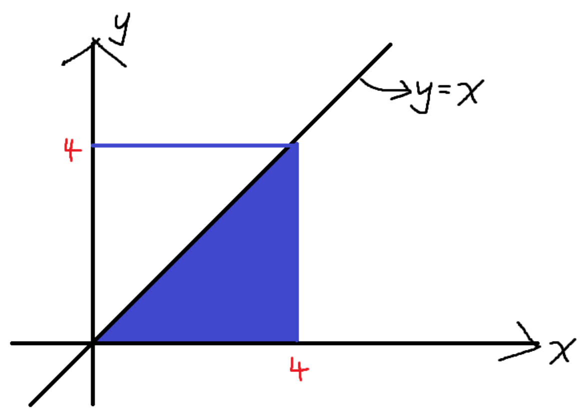


Figure above shows the graph of $y = x$. The blue area is the area we were calculating in the previous example. Indeed, we need not use definite integrals to calculate this area. Notice that the blue area is the area of a triangle with base 4 and height 4. Therefore,

$$\int_0^4 x\,dx = \frac{1}{2}4 \cdot 4 = 8$$

[5]The graph of $y = x$ is actually a straight line. We generalised the notion of curve to include straight lines as well.

Example 4.18 Find $\int_0^2 x^2+1\,dx$.

$$\int_0^2 x^2+1\,dx = \left[\frac{1}{3}x^3+x\right]_0^2 = \left(\frac{2^3}{3}+2\right)-\left(\frac{0^3}{3}+0\right) = \frac{14}{3}$$

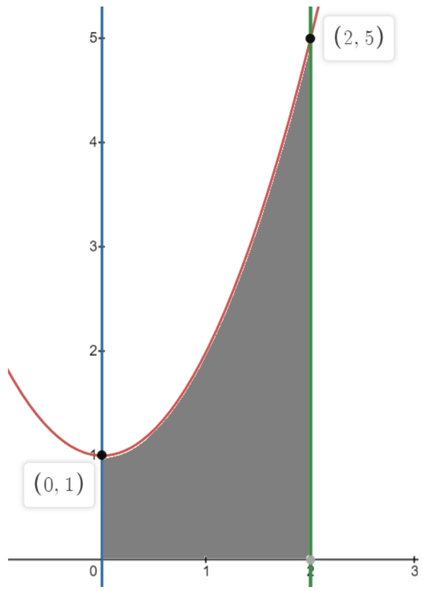


Figure above shows the curve $y=x^2+1$ generated using Desmos. According to the example above, the grey region has area $\frac{14}{3}$ unit^2.

Notice that finding antiderivatives is a crucial step in evaluating definite integrals. We can similarly use the technique of integration by substitution as well as integration by parts, to help us in finding antiderivatives, which in turn helps us to evaluate definite integrals. Note that after each substitution, the bounds of integration needs to be changed accordingly.

Example 4.19 Evaluate $\int_0^2 2x(x^2+4)^2\,dx$.

Let $u=x^2+4$. Then, $du=2x\,dx$.
When $x=0$, $u=4$. When $x=2$, $u=2^2+4=8$. Therefore,

$$\int_0^2 2x(x^2+4)^2\,dx = \int_4^8 u^2\,du = \left[\frac{u^3}{3}\right]_4^8 = \frac{8^3}{3}-\frac{4^3}{3} = \frac{448}{3}$$

When evaluating definite integrals with the help of integration by parts, the bounds of integration does not change, but has to be applied to the uv term in the formula after integration by parts is performed. Informally,

$$\int_a^b u\,dv = [uv]_a^b - \int_a^b v\,du$$

Example 4.20 Evaluate $\displaystyle\int_2^4 x\ln x\,dx$.

Let $u = \ln x$ and $dv = x\,dx$. Then, $du = \dfrac{1}{x}dx$ and $v = \dfrac{x^2}{2}$. Therefore,

$$\begin{aligned}
\int_2^4 x\ln x\,dx &= \left[\frac{x^2}{2}\ln x\right]_2^4 - \int_2^4 \frac{x^2}{2}\frac{1}{x}\,dx \\
&= \frac{4^2}{2}\ln 4 - \frac{2^2}{2}\ln 2 - \int_2^4 \frac{x}{2}\,dx \\
&= 8\ln 4 - 2\ln 2 - \left[\frac{x^2}{4}\right]_2^4 \\
&= 8\ln 4 - 2\ln 2 - \left(\frac{4^2}{4} - \frac{2^2}{4}\right) \\
&= 8\ln 4 - 2\ln 2 - 3
\end{aligned}$$

Bounds of integration may also be ∞ or $-\infty$. In such cases, when the antiderivative has been found, instead of evaluating it at the corresponding bounds, we take its limit as x tends to that bound. In other words, let $F(x)$ be the antiderivative of $f(x)$, then

$$\int_a^\infty f(x)\,dx = \lim_{x\to\infty} F(x) - F(a)$$

$$\int_{-\infty}^b f(x)\,dx = F(b) - \lim_{x\to-\infty} F(x)$$

$$\int_{-\infty}^\infty f(x)\,dx = \lim_{x\to\infty} F(x) - \lim_{x\to-\infty} F(x)$$

Example 4.21 Evaluate $\displaystyle\int_0^\infty e^{-x}\,dx$

The antiderivative of e^{-x} is $-e^{-x}$. Therefore,

$$\int_0^\infty e^{-x}\,dx = [-e^{-x}]_0^\infty = \lim_{x\to\infty}(-e^{-x}) - (-e^0) = 0 - (-1) = 1$$

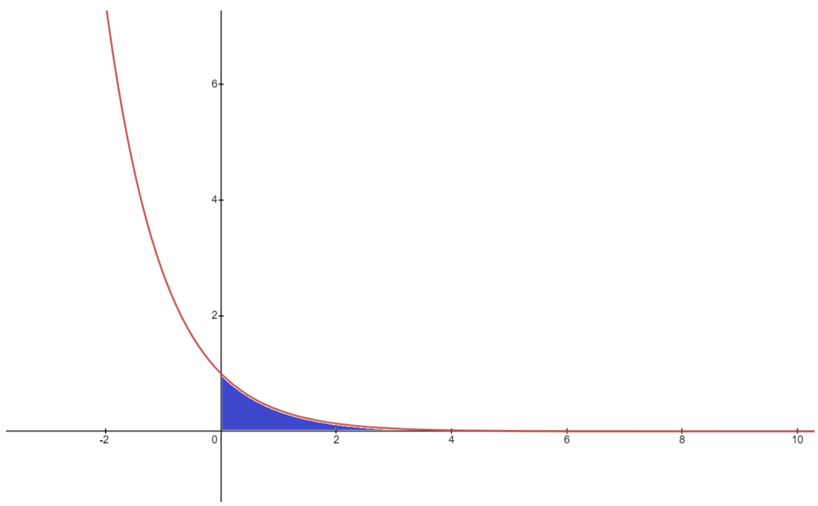


Figure above shows the graph of $y = e^{-x}$ generated using Desmos. In the above example, we found that the area of the blue shaded region is 1 unit^2.

6 Multivariable Differentiation

Multivariable Functions

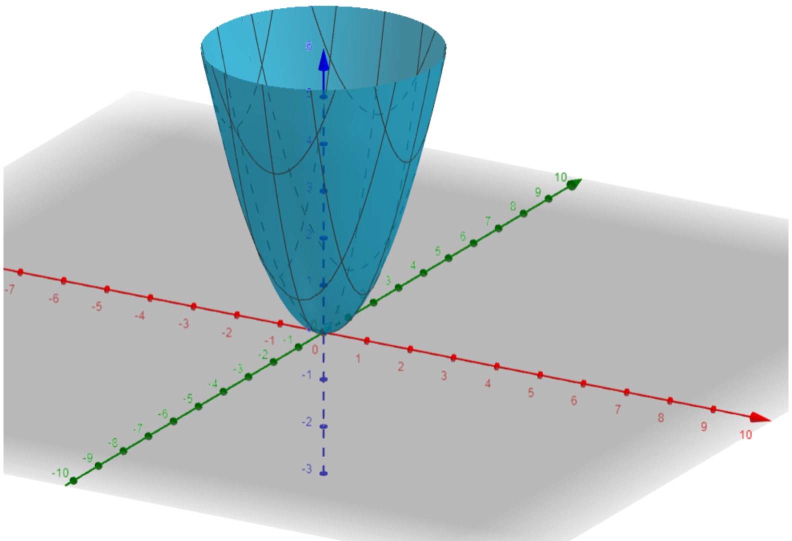
So far, we are only concerned with functions of one variable like $f(x) = x^2$. However, functions can also be multivariable. A **multivariable function** is a function of more than one variables. The following are some examples of multivariable functions:

- $f(x, y) = x^2 y$
- $g(x, y, z) = 2x + 3y + 9z$
- $Q(K, L) = 10K^{0.6}L^{0.4}$
- $U(x_1, x_2, \cdots, x_n) = x_1 + x_2 + \cdots + x_n$

Note that multivariable functions have multiple input values, but still only have one output value. For example, $f(5, -3) = 5^2(-3) = -75$.

For most of the rest of the discussion, we shall focus on functions of two variables, though the concepts presented are also applicable to functions of more than two variables.

We can graph functions of one variable on a 2D plane. Functions of two variables like $f(x, y)$ have to be graphed in a 3D space. Such graphs consist of points (x, y, z) satisfying $f(x, y) = z$. While graphs of functions of one variable tend to be a curve, graphs of functions of two variables tend to be a surface. For example, figure below shows the graph of $z = f(x, y) = x^2 + y^2$, generated using GeoGebra.



In this graph, the red axis represents the x-axis, the green axis represents the y-axis, and the blue axis represents the z-axis.

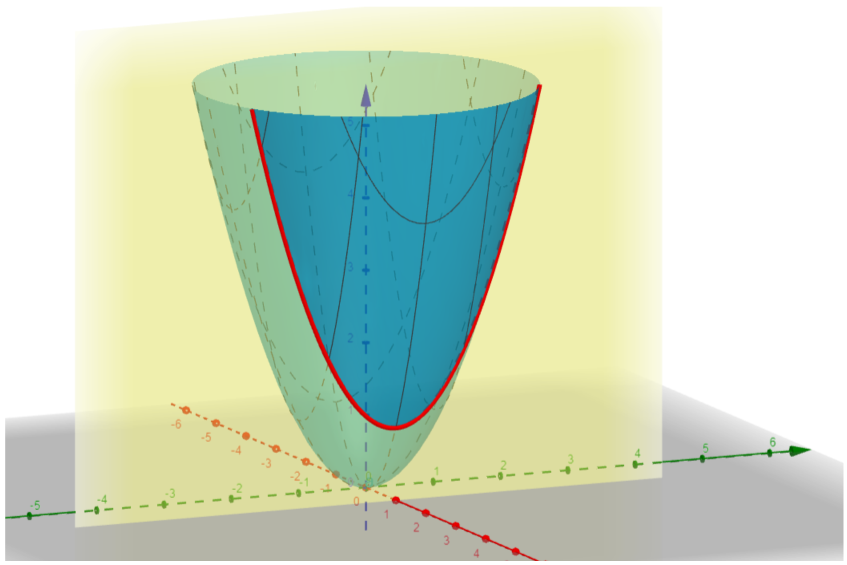
Partial Evaluations

Now, imagine slicing the 3D surface of $f(x, y) = x^2 + y^2$ using a "cardboard" $x = 1$. As shown in the figure in the next page, notice that a 2D curve is formed on the cardboard. The yellow surface represents the cardboard $x = 1$. The red curve is formed on the cardboard, representing the intersection between the cardboard and the surface $x^2 + y^2$.

A very important observation here is that this red curve is the curve of the function $g(y) = f(1, y) = 1 + y^2$.

More generally, the curve obtained by intersecting a cardboard[6] $x = a$ with a surface $f(x, y)$ should be the curve of the function $g(y) = f(a, y)$. Similarly, we can intersect $y = b$ with $f(x, y)$ to obtain the curve of the function $g(x) = f(x, b)$.

[6] The more formal term for a cardboard is a plane.



Given a function $f(x)$, we can evaluate it at $x = a$ to obtain $f(a)$. Here, $f(a)$ is a number. We can also think of it as a new function $g()$[7] of zero variables. There are no variables we can control, hence the output of g is a constant number.

We can similarly evaluate a function $f(x, y)$ at $x = a$ and $y = b$ and obtain a number $f(a, b)$. In addition of this, we previously observed that we can evaluate $f(x, y)$ at $x = a$ only, without specifying the value of y, the result of which is a new function $g(y) = f(a, y)$ of one variable.

In mathematics, we say that $f(x, y)$ is **partially evaluated** at $x = a$[8]. The function is evaluated, but not fully evaluated into a number. In general, when we partially evaluate a function of n variables with m variables, the result will be a new function of $(n - m)$ variables.

Partial Derivatives

Recall how we differentiate functions of one variable. If $f(x) = x^3$, then

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

If we were to differentiate the function $f(y) = y^3$ instead, we may write

$$f'(y) = \frac{d}{dy}y^3 = 3y^2$$

[7]This is not a standard mathematical notation.

[8]In computer programming, this concept is also known as **currying**. This name does not originate from the food, but from the mathematician Haskell Curry.

When it comes to multivariable functions like $f(x,y) = x^2y$, we consider the gradient functions of the function obtained when the functions are partially evaluated. For example, the gradient function of $f(x,2) = 2x^2$ is $4x$. The gradient function of $f(x,3) = 3x^2$ is $6x$. More generally, the gradient function of $f(x,y)$ when y is kept constant is $2xy$.

Definition 5.1 The gradient function of $f(x,y)$ where y is kept constant is known as the **partial derivative** of $f(x,y)$ with respect to x and is denoted by f_x or $\dfrac{\partial}{\partial x}f(x,y)$.

To find the partial derivative of $f(x,y)$ with respect to x, we treat y as a constant and perform differentiation as if f is a function of one variable.

Example 5.2 Given $f(x,y) = 3x^4 + 2x^3y^5 + 4y^7$, find f_x and f_y.

When finding f_x, y^5 and y^7 are treated as constants. Therefore,

$$f_x = 12x^3 + 2 \cdot 3x^2y^5 + 0 = 12x^3 + 6x^2y^5$$

Similarly, treating x^4 and x^3 as constants,

$$f_y = 0 + 2 \cdot 5x^3y^4 + 28y^6 = 10x^3y^4 + 28y^6$$

Example 5.3 Given $f(x,y) = x^2e^{xy}$, find the gradient function of $f(x,1)$.

One method is to find $f(x,1)$, then find the gradient function of $f(x,1)$.

$$f(x,1) = x^2e^x$$

Therefore, the gradient function of $f(x,1)$ is $2xe^x + x^2e^x$.

We may also find f_x, the partial derivative of $f(x,y)$ with respect to x, then evaluate the partial derivative at $y = 1$.

$$f_x = 2xe^{xy} + x^2ye^{xy}$$

Substituting $y = 1$, we obtain the desired gradient function $2xe^x + x^2e^x$.

Example 5.4 Given $z = e^{xy} + \dfrac{x}{y}$, find $\dfrac{\partial z}{\partial x}$ and $\dfrac{\partial z}{\partial y}$.

$$\frac{\partial z}{\partial x} = ye^{xy} + \frac{1}{y}$$

$$\frac{\partial z}{\partial y} = \frac{\partial}{\partial y}(e^{xy} + xy^{-1}) = xe^{xy} - xy^{-2} = xe^{xy} - \frac{x}{y^2}$$

Chain Rule for Partial Derivatives

When y is a function of x and x is a function of t, then f is indirectly a function of t. For example, if $y = \sqrt{x}$ and $x = 4t - 5$, then $y = \sqrt{4t-5}$. We can then apply chain rule to find $y'(t)$ as follows

$$y'(t) = \frac{dy}{dt} = \frac{dy}{dx} \cdot \frac{dx}{dt} = \frac{1}{2}x^{-\frac{1}{2}} \cdot 4 = \frac{2}{\sqrt{x}} = \frac{2}{\sqrt{4t-5}}$$

For a multivariable function $f(x, y)$, if x and y are also functions of t, then f is indirectly a single-variable function of t. To find $\frac{df}{dt}$, we need a similar, more generalised version of chain rule.

Theorem 5.5 Let $f(x(t), y(t))$, then

$$\frac{df}{dt} = \frac{\partial f}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial f}{\partial y} \cdot \frac{dy}{dt}$$

Example 5.6 Let $f(x, y) = x^2 y$, $x(t) = t^2$ and $y(t) = t^3$. Find $\frac{df}{dt}$.

$$\frac{\partial f}{\partial x} = 2xy$$
$$\frac{\partial f}{\partial y} = x^2$$
$$\frac{dx}{dt} = 2t$$
$$\frac{dy}{dt} = 3t^2$$

Therefore,

$$\begin{aligned} \frac{df}{dt} &= \frac{\partial f}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial f}{\partial y} \cdot \frac{dy}{dt} \\ &= 2xy(2t) + x^2(3t^2) \\ &= 4xyt + 3x^2t^2 \\ &= 4(t^2)(t^3)t + 3(t^2)^2t^2 \\ &= 7t^6 \end{aligned}$$

Note that we can also first substitute for x and y and then differentiate the resulting single-variable function, as follows

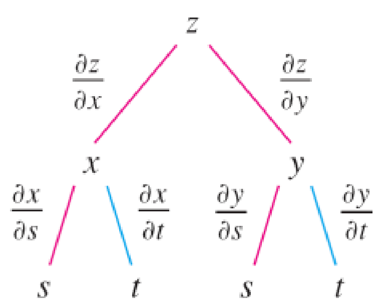
$$f(x, y) = x^2 y = (t^2)^2 t^3 = t^7$$
$$\frac{df}{dt} = 7t^6$$

If f is a function of x and y, and x and y are also multivariable functions of s and t, then f is indirectly a function of s and t. We can again employ a more extended version of chain rule to find $\dfrac{\partial f}{\partial s}$ and $\dfrac{\partial f}{\partial t}$.

Theorem 5.7 Let $f(x(s,t), y(s,t))$, then

$$\frac{\partial f}{\partial s} = \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial s} + \frac{\partial f}{\partial y} \cdot \frac{\partial y}{\partial s}$$
$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial t} + \frac{\partial f}{\partial y} \cdot \frac{\partial y}{\partial t}$$

There is a tree diagram we can draw to help us remember the generalised chain rule.



To find $\dfrac{\partial z}{\partial s}$, we find the product of the partial derivatives along each path from z to s (coloured red) and then add these products. Similarly, paths from z to t give us $\dfrac{\partial z}{\partial t}$.

Total Differential

Partial derivatives help us to understand how does the value of a multivariable change when one of its variables changes by a small amount. On the other hand, if all the variables change by a small amount, we would need to employ the concept of total differential.

Definition 5.8 The **total differential** of a function $f(x,y)$, denoted by df is the change in the value of f when x changes by a small amount dx and y changes by a small amount dy.

Intuitively, if more than one variables are changing, the sum of the effect of change in each variable gives us the overall net change. Recall Theorem 2.18, which states that when a variable x changes by a small amount dx, the corresponding function changes by $f'(x)dx$. Likewise, for a multivariable function

$f(x, y)$, when x changes by a small amount dx and y changes by a small amount dy, then $f(x, y)$ should change by $f_x dx + f_y dy$.

Theorem 5.9 Let $f(x, y)$ be a multivariable function, then

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy$$

Example 5.10 Let $f(x, y) = ye^x$, find the total differential df.

$$\frac{\partial f}{\partial x} = \frac{\partial}{\partial x} ye^x = ye^x$$
$$\frac{\partial f}{\partial y} = \frac{\partial}{\partial y} ye^x = e^x$$

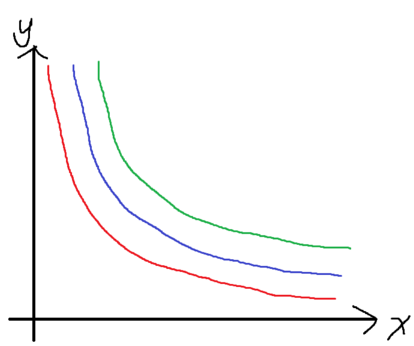
Therefore,

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy = ye^x dx + e^x dy$$

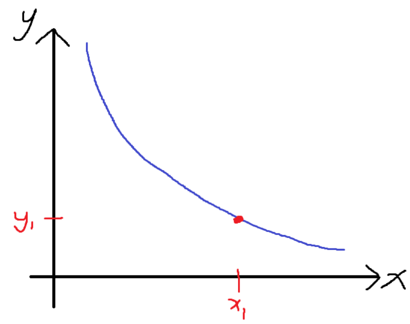
Marginal Rate of Substitution

Given a utility function $U(x, y)$ for some goods x and y, a consumer might give up some amount of goods x in exchange for good y while maintaining the same level of utility, or vice versa. The **marginal rate of substitution (MRS)** of good y for x highlights how many units of y would be considered by the consumer to be compensation for one less unit of x. In other words, if the consumer gives up one unit of x, MRS tells us the amount of units of y required by the consumer to maintain the same level of utility.

If we fix the level of utility U at some constant value, we obtain a relationship between x and y and is able to draw a graph that, most of the time, looks like the following



In the previous figure, each curve of different colour represents the graph of y against x when U is kept at different constants. These curves are known as **indifference curves**.



Suppose that we, as a consumer, wish to maintain our level of utility at a value U such that our indifference curve is the one in blue. We currently have x_1 amount of good x and y_1 amount of good y. Thus, we correspond to the red point in the figure above.

The important observation here is that, if we were to give up some amount of x while maintaining the same level of utility, from the perspective of indifference curve, this is equivalent as moving the red point some units to the left while staying on the blue curve.

If we give up dx amount of x and require dy amount of y to maintain the same level of utility, the MRS is defined to be $\frac{dy}{dx}$.

Now, by chain rule,

$$\frac{dy}{dx} = \frac{\partial y}{\partial U} \cdot \frac{\partial U}{\partial x} = \frac{\frac{\partial U}{\partial x}}{\frac{\partial U}{\partial y}}$$

$\frac{\partial U}{\partial x}$ is known as the marginal utility of x and is denoted by MU_x. $\frac{\partial U}{\partial y}$ is known as the marginal utility of y and is denoted by MU_y.

To summarise, we have

$$\text{MRS} = \frac{dy}{dx} = \frac{MU_x}{MU_y}$$

Note that we have been describing the MRS of good y for x, which means that we are giving up x. If we were to give up y instead, then

$$\text{MRS} = \frac{dx}{dy} = \frac{MU_y}{MU_x}$$

Example 5.11 Let $U(x, y) = x^{0.2}y^{0.8}$, find the MRS of good y for x.

$$MU_x = \frac{\partial U}{\partial x} = 0.2x^{-0.8}y^{0.8}$$
$$MU_y = \frac{\partial U}{\partial y} = 0.8x^{0.2}y^{-0.2}$$

Therefore,

$$\text{MRS} = \frac{MU_x}{MU_y} = \frac{0.2x^{-0.8}y^{0.8}}{0.8x^{0.2}y^{-0.2}} = \frac{1}{4}x^{-1}y = \frac{y}{4x}$$