Quantitative Methods in Economics
ECON 271:10
Chapter 3
Derivatives
& their Applications in Economics

Teng Wah Leo

We will now be getting into derivatives, which is at the heart of Calculus. We begin by building an intuition through examining its brief development through common problems in Mathematics. After that, we will introduce the Economics component.

1 Some Examples to Motivate the Idea of Derivatives

Consider first the problem of finding the slope of a curve at a specific point.

Problem 1 ***Slope.*** *What is the slope of a tangent line, say to $y = x^2$, at a point $P = (1, 2)$? By tangent line here, we will first adopt a rudimentary definition that means a line through P that has the "same direction" as the curve at P.*

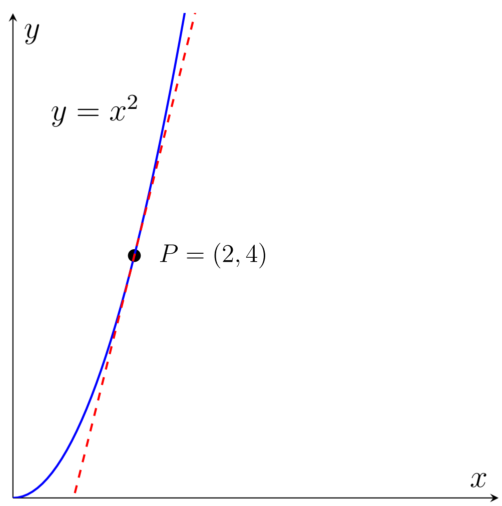
Solution 1 *The immediate problem you face is how you calculate a slope with only one point given, the diagram of which is in figure 1? Let us choose another point on $y = x^2$ that is very close to P, and use that. Let $Q = (2.1, 2.1^2)$, where the line $P - Q$ is known as a secant. Then the slope through points P and Q is $\frac{2.1^2-2^2}{2.1-2} = 4.1$. However, this* *is merely an estimate, since choosing another alternative* Q, $Q' = (2.01, 2.01^2)$, *yields another estimate. In deed, we can take this idea, and formalize the calculation into a formula of sorts,*

$$\frac{(2+h)^2 - 2}{h}$$

Can we find the actual slope at P *of* $y = x^2$ *using what we have learned in limits, by bringing* $h \to 0$*?*

$$\begin{aligned}\lim_{h\to 0}\frac{(2+h)^2-2^2}{h} &= \lim_{h\to 0}\frac{4+4h+h^2-4}{h}\\ &= \lim_{h\to 0}(4+h) = 4\end{aligned}$$

Figure 1: Slope of $y = x^2$ at $P = (1, 2)$



Lets consider another problem, now to do with speed.

Problem 2 ***Speed.*** *Consider a stone at rest falls* $16t^2$ *feet in* t *seconds. What then is the speed achieved after 2 seconds?*

Solution 2 *What the problem as asking is instantaneous speed, or "speed at a given time". The closest most of us can concieve of as a formula is the distance travelled per unit time, here per second, which means* $\frac{\text{distance}}{\text{time}}$*. So given the information, we can say that at the start of the 2 second mark, the stone would have travelled 64 feet. How about at the 2.01 second mark? 64.6416 feet. So over the split 0.1 second, the average speed is* $\frac{0.6416}{0.01} = 64.16$ *feet per second. We can again reformulate this in terms of small amounts* *of change* h*, in time now. So we have change in distance as* $16(2+h)^2 - 16.2^2$ *feet, giving us,*

$$\begin{aligned}\lim_{h\to 0} \frac{16(2+h)^2 - 16.2^2}{h} &= 16 \lim_{h\to 0} \frac{(2+h)^2 - 2^2}{h} \\ &= 16.4 = 64 \text{ ft/sec}\end{aligned}$$

You may now observe that in both examples, we have a non-linear, or curved "trajectory" that we deal with, and that using the idea of limits in combination with continuity of a function allows us to precisely state a "rate of change" measure that goes beyond what you known in straight trajectories. In deed, the idea may be expanded into considerations of magnification, density, volume changes ....

2 The Derivative

In the prior discussion, in forming an expression so as to permit the use of limits, we created a *difference quotient*,

$$\frac{\text{Difference in Output}}{\text{Difference in Inputs}} = \frac{\text{Change in } y}{\text{Change in } x}$$

before examining its limits. This underlying mathematical concept is known as a *derivative*, and is defined as follows.

Definition 1 ***The derivative of a function at*** x. *Let* f *ne a function that is defined at least in some open interval that includes* x*. If*

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

exists, it is called the ***derivative of*** f ***at*** x *and is commonly denoted as* $\frac{df(x)}{dx} = \frac{dy}{dx} = f'(x)$*. The function is said thus to be* ***differentiable*** *at* x*.*

If the function f instead is defined only to the right of x in the interval $[x, b)$, then in the definition above, the limit would be replaced by $h \to 0^+$. If instead the function is defined to the left of x on the interval $(a, x]$, the limit would instead be $h \to 0^-$.

With this, we can start seeing a pattern in derivatives, so that formulas may be discovered.

Example 3 *Find* $f'(x)$ *for* $y = f(x) = x^3$*.*

Solution 3 *Utilizing the same ideas prior, we can formulate the limit of this function for a small arbitrary change in x by h,*

$$\begin{aligned} f'(x) = \lim_{h\to 0} \frac{(x+h)^3 - x^3}{h} &= lim_{h\to 0} \frac{x^3 + 3x^2h + 3xh^2 + h^3 - x^3}{h} \\ &= \lim_{h\to 0} \frac{3x^2h + 3xh^2 + h^3}{h} \\ &= \lim_{h\to 0} (3x^2 + 3xh + h^2) = 3x^2 \end{aligned}$$

Let's consider instead an exponent that is $(0, 1)$.

Example 4 *Find the derivative of $f(x) = \sqrt{x}$.*

Solution 4 *Using the same ideas as the previous example,*

$$\begin{aligned} f'(x) &= \lim_{h\to 0} \frac{\sqrt{(x+h)} - \sqrt{x}}{h} \\ &= \lim_{h\to 0} \frac{\sqrt{(x+h)} - \sqrt{x}}{h} \times \left(\frac{\sqrt{(x+h)} + \sqrt{x}}{\sqrt{(x+h)} + \sqrt{x}} \right) \\ &= \lim_{h\to 0} \frac{x + h - x}{h(\sqrt{(x+h)} + \sqrt{x})} \\ &= \lim_{h\to 0} \frac{1}{\sqrt{(x+h)} + \sqrt{x}} = \frac{1}{2\sqrt{x}} \end{aligned}$$

Note that at $x = 0$ the limit does not exist, so that the derivative of $\sqrt{x}$ does not exist at $x = 0$.

You can keep trying various polynomials, but what you will arrive at is the following theorem.

Theorem 1 *For each positive integer n, for $y = f(x) = x^n$*

$$\frac{dy}{dx} = \frac{dx^n}{dx} = x^{n\prime} = nx^{n-1}$$

Proof.

$$\begin{aligned} \frac{(x+h)^n - x^n}{h} &= \frac{x^n + nx^{n-1}h + g(h^2) - x^n}{h} \\ &= nx^{n-1} + g'(h) \end{aligned}$$

where $g(h^2)$ is some function of h^2 terms and $g'(h) = \frac{g(h^2)}{h}$, since we know as we expand on $(x+h)^n$ we get terms that have h with increasing exponents upto n as revealed by our examination of the derivative of x^3. This means that those terms after dividing by h, would still be multiplied by at least h with at least 1 as the exponent, so that as we take limits $h \to 0$, $g(h^2) \to 0$ and $g'(h) \to 0$. Therefore,

$$\begin{aligned}
\lim_{h\to 0} \frac{(x+h)^n - x^n}{h} &= \lim_{h\to 0} \left(\frac{x^n + nx^{n-1}h + g(h^2) - x^n}{h} \right) \\
&= \lim_{h\to 0} \left(nx^{n-1} + g'(h) \right) = nx^{n-1}
\end{aligned}$$

as needed. ■

The next theorem addresses the example on $\sqrt{x}$.

Theorem 2 *For any positive integer n, for $f(x) = x^{\frac{1}{n}}$,*

$$\frac{dx^{\frac{1}{n}}}{dx} = \frac{1}{n} x^{\frac{1}{n}-1}$$

for which both $x^{\frac{1}{n}}$ and $x^{\frac{1}{n}-1}$ are defined.

Proof. For this proof, we would need to make use of the algebraic identity,

$$d^n - c^n = (d-c)(d^{n-1} + d^{n-2}c + \cdots + dc^{n-2} + c^{n-1})$$

Let $d = (x+h)^{1/n}$ and $c = x^{1/n}$, then examining the limit we need for the theorem,

$$\begin{aligned}
&\lim_{h\to 0} \frac{(x+h)^{1/n} - x^{1/n}}{h} \\
= &\lim_{h\to 0} \frac{(x+h)^{1/n} - x^{1/n}}{h} \times \frac{[(x+h)^{1/n}]^{n-1} + [(x+h)^{1/n}]^{n-2}x^{1/n} + \cdots + (x^{1/n})^{n-1}}{[(x+h)^{1/n}]^{n-1} + [(x+h)^{1/n}]^{n-2}x^{1/n} + \cdots + (x^{1/n})^{n-1}} \\
= &\lim_{h\to 0} \frac{[(x+h)^{1/n}]^n - (x^{1/n})^n}{h} \frac{1}{[(x+h)^{1/n}]^{n-1} + [(x+h)^{1/n}]^{n-2}x^{1/n} + \cdots + (x^{1/n})^{n-1}} \\
= &\lim_{h\to 0} \frac{x+h-x}{h} \frac{1}{[(x+h)^{1/n}]^{n-1} + [(x+h)^{1/n}]^{n-2}x^{1/n} + \cdots + (x^{1/n})^{n-1}} \\
= &\frac{1}{(x^{1/n})^{n-1} + \cdots + (x^{1/n})^{n-1}} \\
= &\frac{1}{nx^{\frac{n-1}{n}}} = \frac{1}{nx^{1-\frac{1}{n}}} = \frac{1}{n} x^{\frac{1}{n}-1}
\end{aligned}$$

■

2.1 Derivatives & Continuity; Antiderivatives

The above two theorems from the previous section are collectively commonly known as the power function rule. We can derive additional common rule that constitute the *rules of* *differentiation* that provide the formulas to obtain the derivatives to a variety of functions. But before we do so, we will introduce additional additional notation, the relationship between derivatives and continuity, and introduce the idea of an *antiderivative*.

It is common to write the difference between $(x + h)$ and x as Δx, so that we can define a derivative as,

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

Further the difference in the functions in the numerator are commonly written as Δf, so that the definition may be further refined as,

$$f'(x) = \lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x}$$

This latter expression explains why we typically write $f'(x)$ as $\frac{dy}{dx}$ or $\frac{df}{dx} \equiv \frac{df(x)}{dx}$ since Δ is d in greek. You may also see in articles $Dy \equiv Df \equiv Df(x)$, which all mean the same thing, the derivative of the function f with respect to x.

As you progress along in economics or mathematics or statistics or their various subfields, you may observe the notation $\dot{x}$. This is the notation used by Newton, and refers to $\frac{dx}{dt}$, where t is time. In consequence, you will see that notation when and if you start dealing with dynamics, such as in Advanced Macroeconomics.

You would have observed, based on figure 1 which illustrates the genesis of the idea of differentiation, that the function itself is continuous, as we have defined it. This is indeed necessary, in other words, you can take a derivative only for continuous functions. Put another way, if a function f is differentiable at each value of x on some interval, we then say that f is differentiable throughtout that interval. The following theorem tells you that a differentiable function must be continuous.

Theorem 3 *If f is differentiable at a, then it is continuous at a.*

Proof. The sequence of proof is instructive, and highlights the importance of sequential reasoning.

We need to show first that $\lim_{x \to a} f(x)$ exists, and is equal to $f(a)$. This is because if the limit exists, and is continuous at a, the two expression must be the same. Thus the two ideas can meld into one expression, to show

$$\lim_{x \to a} (f(x) - f(a)) = 0$$

is true, because then

$$\begin{aligned}\lim_{x\to a} f(x) &= \lim_{x\to a}[(f(x)-f(a))+f(x)] \\ &= 0+f(a)=f(a)\end{aligned}$$

Thus, to begin with, for $x \neq a$,

$$\begin{aligned}f(x)-f(a) &= \frac{f(x)-f(a)}{x-a}(x-a) \\ \text{so that } \lim_{x\to a}(f(x)-f(a)) &= \lim_{x\to a}\frac{f(x)-f(a)}{x-a}(x-a) \\ &= \lim_{x\to a}\frac{f(x)-f(a)}{x-a}.\lim_{x\to a}(x-a) \text{ (since both limits exists)} \\ &= f'(a).0=0\end{aligned}$$

So that f is continuous at a. Note that the differentiability of f is not in question, but the continuity of f at a is. ■

Finally, to complete the discussion here, for two functions f and F, where $f = DF$, in other words, f is the derivative of F, we call F then the *antiderivative* of f. This idea has got to do with the "reverse" process of differentiation, known as *integral.*

2.2 Rules of Differentiation & Comparative Statics

Differentiation or finding derivatives of a function can be quite routine, once the key rules are understood. We will now discuss them in some detail.

Theorem 4 ***Constant Function Rule.*** *For a function $f(x) = c$ where c is a constant,*

$$\frac{df}{dx}=0$$

Proof.

$$\begin{aligned}\frac{df}{dc} &= \lim_{\Delta x\to 0}\frac{f(x+\Delta x)-f(x)}{\Delta x} \\ &= \lim_{\Delta x\to 0}\frac{c-c}{\Delta x}=0\end{aligned}$$

■

This is theorem is quite intuitive. If the function is a horizontal line, then there is no slope or change to speak of as we change the independent variable or input x.

The next theorem is likewise quite intuitive. It provides the rule for differentiating the sum or difference of two differentiable functions.

Theorem 5 *If f and g are differentiable functions, so too are $f+g$ and $f-g$, and their derivatives are given by,*

$$\frac{d(f \pm g)}{dx} \quad = \quad \frac{df}{dx} \pm \frac{dg}{dx}$$

Proof.

$$\begin{aligned}
\frac{d(f \pm g)}{dx} &= \lim_{\Delta x \to 0} \left(\frac{[f(x+\Delta x) \pm g(x+\Delta x)] - [f(x) \pm g(x)]}{\Delta x} \right) \\
&= \lim_{\Delta x \to 0} \left(\frac{[f(x+\Delta x) - f(x)] \pm [g(x+\Delta x) - g(x)]}{\Delta x} \right) \\
&= \lim_{\Delta x \to 0} \left(\frac{f(x+\Delta x) - f(x)}{\Delta x} \pm \frac{g(x+\Delta x) - g(x)}{\Delta x} \right) \\
&= \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} \pm \lim_{\Delta x \to 0} \frac{g(x+\Delta x) - g(x)}{\Delta x} \\
&= \frac{df}{dx} \pm \frac{dg}{dx}
\end{aligned}$$

■

Although the theorem is couched in terms of two functions only, this idea holds for larger numbers of functions. That is generally, for differentiable functions $f_1 \ldots f_k$ where each function is indexed by the subscript,

$$\frac{d(f_1 \pm \cdots \pm f_k)}{dx} = \frac{df_1}{dx} \pm \cdots \pm \frac{df_k}{dx}$$

This next theorem pertains to product of two functions, and can likewise be generalized to the product of a large number of functions.

Theorem 6 ***Product Rule.*** *If f and g are differentiable functions, so too are $f.g$, and their derivative is,*

$$\frac{d(f.g)}{dx} = g\frac{df}{dx} + f\frac{dg}{dx}$$

Proof. For the proof, we will use a simpler notation we had used earlier.

$$\begin{aligned}
\Delta f &= f(x+\Delta x) - f(x) \\
\Rightarrow f(x+\Delta x) &= \Delta f + f(x) \\
\text{or simply as } f(x+\Delta x) &= \Delta f + f
\end{aligned}$$

Likewise, we can write $g(x+\Delta x) = \Delta g + g$

So to begin,

$$\begin{aligned}\Delta(f.g) &= f(x+\Delta x)g(x+\Delta x) - f(x)g(x) \\ &= (\Delta f + f)(\Delta g + g) - f(x)g(x) \\ &= \Delta f \Delta g + g\Delta f + f\Delta g + f(x)g(x) - f(x)g(x)\end{aligned}$$

Therefore,

$$\begin{aligned}\lim_{\Delta x \to 0} \frac{\Delta f g}{\Delta x} &= \lim_{\Delta x \to 0} \left(\Delta f \frac{\Delta g}{\Delta x} + g \frac{\Delta f}{\Delta x} + f \frac{\Delta g}{\Delta x} \right) \\ &= g\frac{df}{dx} + f\frac{dg}{dx}\end{aligned}$$

This is because,

$$\lim_{\Delta x \to 0} \Delta f = \lim_{\Delta x \to 0} f(x+\Delta x) - f(x) = 0$$

and the same is true of Δg in $\Delta f.\Delta g$. ■

The following is a corollary to the basic definition of a derivative.

Theorem 7 *If c is a constant function, and f is a differentiable function, then $c.f$ is differentiable too, and its derivative is*

$$\frac{d(cf)}{dx} = c\frac{df}{dx}$$

The proof of the above is using the product rule in concert with the constant function rule, and is left to you.

The next rule is very important, and commonly used, just like the product rule, and is as follows.

Theorem 8 ***Quotient Rule.*** *If functions f and g are differentiable, then so too is $\frac{f}{g}$, and the derivative is,*

$$\frac{d(f/g)}{dx} = \frac{g\frac{df}{dx} - f\frac{dg}{dx}}{g^2}$$

where $g(x) \neq 0$.

Proof.

$$\begin{aligned}
\lim_{\Delta x \to 0} \frac{\Delta(f/g)}{\Delta x} &= \lim_{\Delta x \to 0} \frac{\frac{f(x+\Delta x)}{g(x+\Delta x)} - \frac{f(x)}{g(x)}}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{\frac{g(x)f(x+\Delta x) - f(x)g(x+\Delta x)}{g(x+\Delta x)g(x))}}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{\frac{g[\Delta f + f] - f(\Delta g + g)}{g(x+\Delta x)g(x)}}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{\frac{g\Delta f + f.g - f\Delta g - f.g}{g(x+\Delta x)g(x)}}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{g\frac{\Delta f}{\Delta x} - f\frac{\Delta g}{\Delta x}}{g(x+\Delta x)g(x)} \\
&= \frac{g\frac{df}{dx} - f\frac{dg}{dx}}{g^2}
\end{aligned}$$

■

There are two additional corollary which we will state without proof, which is left to you.

Corollary 1

$$\frac{d(1/g(x))}{dx} = -\frac{\frac{dg}{dx}}{g^2}$$

Corollary 2 *Let n be negative integers, then*

$$\frac{dx^n}{dx} = nx^{n-1}$$

What if functions are are composite functions, in the sense of being composed of several functions say $h = f \circ g$? Is h differentiable if f and g are? If so, how do you calculate it? To get a better impression of what we mean by a composite function, an example would be $y = f(u)$ and $u = g(x)$, so that $h(x) = f(g(x)) \equiv f \circ g(x)$. It turns out that the calculation is not very difficult, and the analogy to its solution is like "pealing an onion", albeit without the tears, with practice. In the example, we can use what is known as the *Chain Rule* to find the derivative,

$$\frac{d(f \circ g)(x)}{dx} = \frac{df}{du}.\frac{du}{dx}$$

Formally stated,

Theorem 9 ***The Chain Rule.*** *If f and g are differentiable functions, then $\forall x$ such that $g(x)$ is in the domain of f, $h = f \circ g$ will also be differentiable at x, and*

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

Proof. Without loss of generality, let $y = f(u)$ and $u = g(x)$, and $g'(x) \neq 0$, then

$$\begin{aligned}
\frac{dh}{dx} &= \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta u}\frac{\Delta u}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta u} \lim_{\Delta x \to 0} \frac{\Delta u}{\Delta x} \\
&= \lim_{\Delta u \to 0} \frac{\Delta y}{\Delta u} \lim_{\Delta x \to 0} \frac{\Delta u}{\Delta x}\text{, since as } \Delta x \to 0,\ \Delta u = g(x + \Delta x) - g(x) = 0 \\
&= \frac{dy}{du}\frac{du}{dx} \\
&= f'(g(u)).g'(x)
\end{aligned}$$

(What is $g'(x) = 0$?) ■ With this, you can prove what you might have surmised, about the power function rule.

Theorem 10 *For rational number r, $\frac{dx^r}{dx} = rx^{r-1}$.*

Proof. For $f(x) = x^r$, we can express r as a ratio of two intergers m and n, m/n for positive n, so that $f(x) = x^{m/n} = \left(x^{1/n}\right)^m$. Let $y = u^m$ and $u = x^{1/n}$. Then applying the Chain Rule,

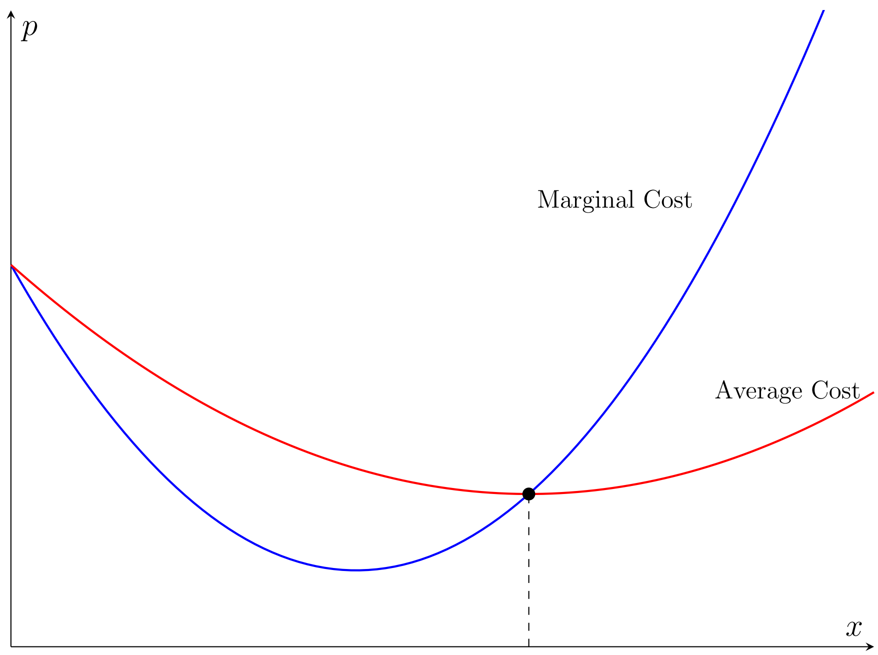
$$\begin{aligned}
\frac{df(x)}{dx} &= mu^{m-1}.\frac{1}{n}x^{\frac{1}{n}-1} \\
&= \frac{m}{n}\left(x^{\frac{m-1}{n}}\right)\left(x^{\frac{1}{n}-1}\right) \\
&= \frac{m}{n}x^{\frac{m-1}{n}+\frac{1}{n}-1} \\
&= \frac{m}{n}x^{\frac{m}{n}-1}
\end{aligned}$$

■

We can keep going, and discover the derivatives for trigonometric functions, but that will be left to a in depth course with the math department should you wish to pursue math further.

To provide an idea of how we can use these skills, we will now examine specific example. In microeconomic discussion of firm choice, we typically depict the marginal cost curve of

Figure 2: The Relationship between a firm's Marginal versus Average Cost



the firm intersecting the average cost from the bottom at its minimum point as seen in figure 2. This idea can be proven using calculus.

Example 5 *Consider a firm with a total cost equation of* $TC = q^3 - 12q^2 + 60q$*. Show that the marginal cost curve intersects the average cost curve at the latter's minimum point.*

Solution 5 *By definition, the average cost is just* $AC = \frac{TC}{q} = q^2 - 12q + 60$*. It minimum point occurs at the point where it has a gradient of 0 based on the diagram 2, which means it occurs at*

$$\begin{aligned} \frac{d(AC)}{dq} &= 2q - 12 = 0 \\ \rightarrow q &= 6 \end{aligned}$$

which gives AC at 24. Marginal cost equation is just,

$$\frac{d(TC)}{dq} = 3q^2 - 24q + 60$$

which at $q = 6$ *achieves an* $MC = AC = 24$*. You may verify that for* $q < 6$*,* $MC < AC$*, while for* $q > 6$*, the* $MC > AC$*, as depicted in figure 2.*

2.3 Partial Differentiation

Thus far, we have dealt with functions with a single independent variable, or input variable. How do we deal with functions that have more than one variable, such as $y = f(x_1, x_2, \ldots, x_k)$? It turns out we can examine the change in y or f by holding the effect of all the other variables unchanged, in other words,

Definition 2 ***Partial Derivatives.*** *If the domain of f includes the region within some sphere around the point $(x_1, \ldots, x_i, \ldots, x_k)$ and*

$$\frac{df(x_1, x_2, \ldots, x_k)}{dx_i} = \lim_{\Delta x_i \to 0} \frac{f(x_1, \ldots, x_{i-1}, x_i + \Delta x_i, x_{i+1}, \ldots, x_k)}{\Delta x_i}$$

exists, this limit is known as a ***partial derivative of f with respect to x_i at*** *$(x_1, \ldots, x_i, \ldots, x_k)$. It is commonly denoted at follows,*

$$f_{x_i} = f_1' = \frac{\partial f(x_1, \ldots, x_i, \ldots, x_k)}{\partial x_i}$$

Note the notation ∂ is known as *partial*, and notice its difference from the greek δ. Practically, what this means is that when finding the derivative with respect to x_i, you would treat x_j, $j \neq i$ as a constant.

If you have already taken microeconomics I you would have been taught the idea that in economics, we think of an individual's wellbeing as being measured by a abstract function known as a *utility function* is a dependent on the quantity of goods he/she consumes. The following example shows how we find the marginal gain in wellbeing from consuming an additional unit of a good.

Example 6 *A consumer has a utility function $u(x_1, x_2) = x_1^2 x_2^3$ constituted by the quantities consumed of two goods, x_1 and x_2. What is the marginal utility from consuming good 2?*

Solution 6 *Given the utility function, the marginal utility is simply,*

$$\begin{aligned} MU_2 &= \frac{\partial u(x_1, x_2)}{\partial x_2} \\ &= x_1^2(2x_2^2) = 2x_1^2 x_2^2 \end{aligned}$$

Consider another example in microeconomics, now from the perspective of a firm.

Example 7 *A firm has a production function* $f(q_1, q_2) = q_1^2 q_2 + q_1 q_2^3$*, in other words, it produces two outputs. What is the marginal product of producing an additional unit of good 1?*

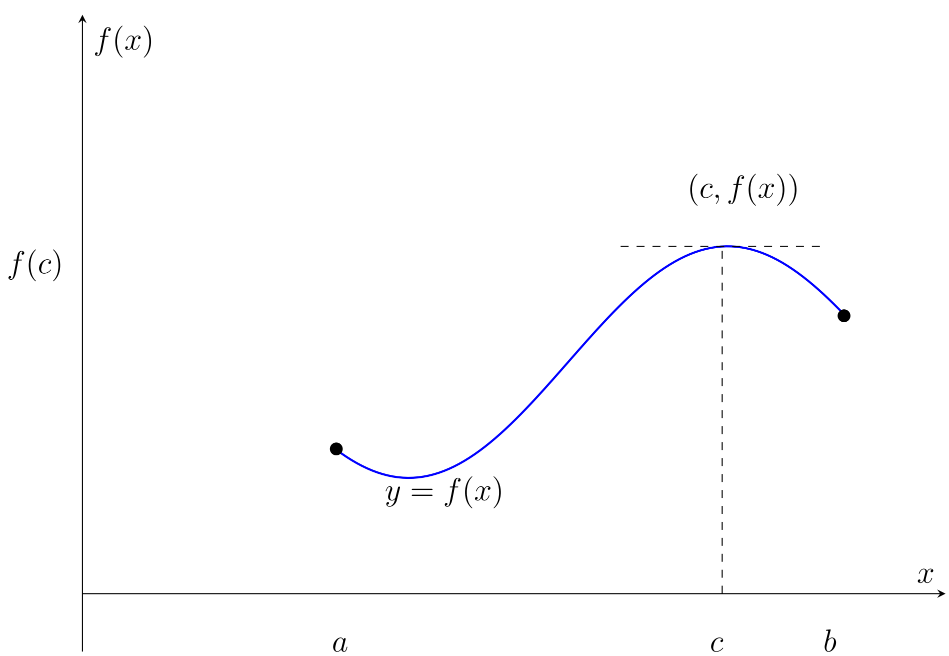
Solution 7 *Given the production function, the marginal product is just,*

$$\begin{aligned} MP_1 &= \frac{\partial f(q_1, q_2)}{\partial q_1} \\ &= 2q_1 q_2 + q_2^3 = \left(2q_1 + q_2^2\right) q_2 \end{aligned}$$

2.4 Rolle's Theorem and the Mean-Value Theorem

Underlying much of the applications of calculus are two important theorems which we will examine now. Suppose function f is defined on an interval $[a, b]$. We know that if it is differentiable, f must be continuous. We also know that based on Maximum value theorem that f must have a maximum at some $c \in [a, b]$, so that $f(c) \geq f(x)\ \forall x \in [a, b]$. What is the implication then of f' at c? As in figure 3, this would mean that at c in the open interval of (a, b), f achieves its maximum. What that means diagrammatically is that the tangent line at $(c, f(c))$ would be parallel to the $x-$axis, so that $f'(c) = 0$. On the other hand, if the extremum occurs at a or b, f' need not be 0.

Figure 3: Interior Extremum



Theorem 11 ***Theorem of the Interior Extremum.*** *For a function f defined on an open interval (a, b), if f achieves an extreme value at c, $c \in (a, b)$, and f' exists, then $f'(c) = 0$.*

Proof. Let $f(c)$ be the maximum value of f on $[a, b]$, and that $f'(c)$ exists. To show that $f'(c) = 0$, we will show in turn $f'(c) \leq 0$ and $f'(c) \geq 0$, so that for these two inequalities to be true, $f'(c) = 0$.

Consider

$$f'(c) = \lim_{\Delta x \to 0} \frac{f(c + \Delta x) - f(c)}{\Delta x}$$

Through it all, keep Δx sufficiently small so that $c + \Delta x \in [a, b]$. Since f achieves its maximum at c, it must be that $f(c) \geq f(c + \Delta x)$, which in turn implies that $f(c + \Delta x) - f(c) \leq 0$. We know Δx can take on either positive or negative values, so we will explore their implications in turn.

When Δx is positive, as $\Delta x \to 0^+$, $\frac{f(c+\Delta x) - f(c)}{\Delta x}$ cannot be positive since the numerator is non-positive (negative or zero), while the denominator is positive. Therefore,

$$f'(c) = \lim_{\Delta x \to 0^+} \frac{f(c + \Delta x) - f(c)}{\Delta x} \leq 0$$

On the other hand, when Δx is negative, as $\Delta x \to 0^-$, $\frac{f(c+\Delta x) - f(c)}{\Delta x}$ cannot be negative since the numerator is non-positive (negative or zero), and the denominator is negative. Therefore,

$$f'(c) = \lim_{\Delta x \to 0^-} \frac{f(c + \Delta x) - f(c)}{\Delta x} \geq 0$$

Since $0 \leq f'(c) \leq 0$, it must be that $f'(c) = 0$. To complete the proof, you should do the same for a minimum as well. ■

This theorem will often be used in mathematics, statistics, and economics to find the maximum or minimum values of a function, since this is often where choices made by agents in the economy occur. Think of an economic agent maximizing her wellbeing, or a firm maximizing its profit. Further, the idea of an extremum does not mean only a maximum, but a minimum as well, as the case may be. This then leads to *Rolle's Theorem.*

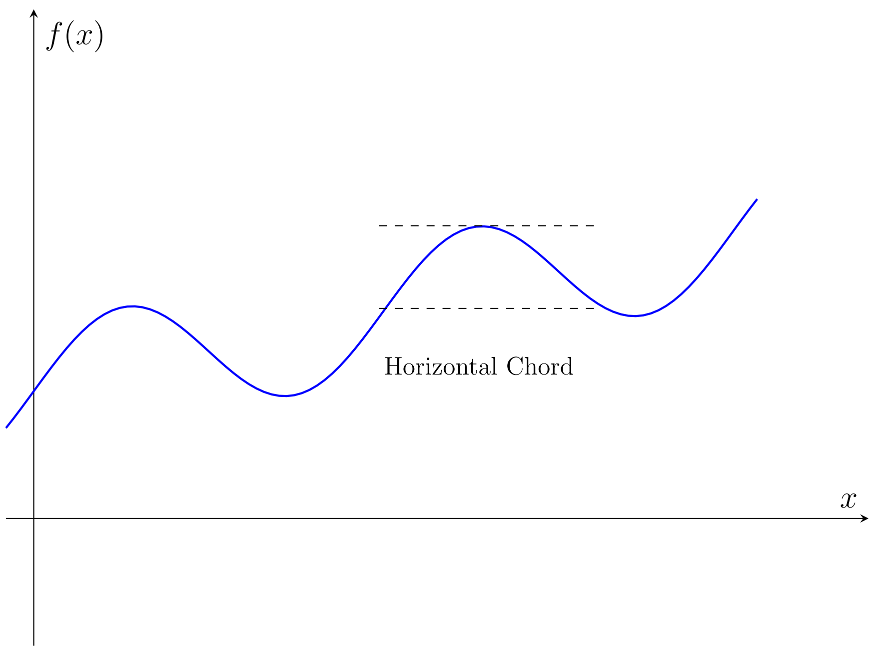
Theorem 12 ***Rolle's Theorem.*** *For a continuous function f on a closed interval $[a, b]$, let f' be defined on an open interval (a, b). If $f(a) = f(b)$, then there must be at least one number $c \in (a, b)$ such that $f'(c) = 0$.*

Proof. We know that for continuous f, it will have a maximum mx and a minimum mn for some $x \in [a, b]$. This necessarily mean that $mx \geq mn$. If $mx = mn$, this would mean that f is constant, a horizontal line, and $f'(x) = 0 \ \forall x \in [a, b]$, so that any $x \in (a, b)$ can be c.

If instead, $mx > mn$, then mx and mn cannot both be at the ends of the domain since $f(a) = f(b)$. This thus mean that at least one of them, mx or mn must be such that $a < c < b$, so that by the interior extremum theorem, $f'(c) = 0$. ■

Define a *chord* as a line segment connecting two points on a graph of f. Rolle's theorem thus says that if a horizontal chord may be drawn for the function, there must be a tangent line to the function that is parallel to the chord much like in figure 4

Figure 4: Rolle's Theorem



In turn, Rolle's Theorem is generalized by the *Mean-Value Theorem* to chords of any angle.

Theorem 13 *For a continuous function f on a closed interval $[a, b]$, differentiable $\forall x \in (a, b)$, there is at least one $c \in (a, b)$ such that,*

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

Proof. Since this theorem generalizes the Rolle's Theorem, we can use a function that permits us to utilize its implications. Define the chord through $(a, f(a))$ and $(b, f(b))$ as part of the line l with a function $g(x)$. Next, let the difference between f and g be $h(x) = f(x) - g(x)$, in other words, the difference betweenn these two graphs. By definition, these differences are the same at the ends of the domain, $h(a) = h(b)$, where they meet.

Then by Rolle's Theorem we know that there must be a c such that $h'(c) = 0$, for $c \in (a, b)$. By our definition $h'(c) = f'(c) - g'(c)$. Since $g(x)$ is just a straight line, we likewise know that its slope is $g'(x) = \frac{f(b)-f(a)}{b-a}$ $\forall x \in (a, b)$. Therefore,

$$\begin{aligned} 0 = h'(c) &= f'(c) - \frac{f(b) - f(a)}{b - a} \\ \Rightarrow f'(c) &= \frac{f(b) - f(a)}{b - a} \end{aligned}$$

■

2.5 Using Derivatives in Solving Problems & Higher Derivatives and Curvature

You might have intuited by the above discussions what calculus can be used for. Not only can they help us in visualizes the shape of functions we work with, but if those functions represent the objectives of the agents within an economy, it could permit us to find the optimal values they would reasonably choose. Before we begin to discuss how this ideas are used, we need some additional definitions. First, when is a point an *important* point.

Definition 3 ***Critical number and critical point.*** *The number c where $f'(c) = 0$ is known as a critical number for function f, and that point $(c, f(c))$ is thus known as a critical point on that graph of f.*

Are all extremums the same in terms of their importance.

Definition 4 ***Relative maximum (local maximum).*** *A function has a relative maximum (local maximum) at a point c if $f(c) \geq f(x)$ $\forall x \in (a, b)$ where (a, b) constitute a portion of the domain of f. We can similarly define a* ***relative minimum (local minimum)****.*

In turn,

Definition 5 ***Global Maximum (absolute maximum).*** *A function f has a global maximum (absolute maximum) at c if $f(c) \geq f(x)\ \forall x$ in the entire domain of f. We can similarly define a* ***global minimum (absolute minimum)****.*

With this, we can think about a sequence of things we can do to find a point associated with a local extremum. For us to determine if there is a local maximum at c, where $c \in (a, b)$ which is in the domain of f. For a continuous f on (a, b), if f is also differentiable on (a, b), except possibly at c, and we find $f'(x) \geq 0\ \forall x < c$, and $f'(x) \leq 0\ \forall x > c$, then f has a local maximum at c. We can reverse the sign to find a local minimum.

However, this seems like it is alot of work. You might wish to make the process quicker. This is where higher order derivatives become useful. What do higher derivatives do? If we look at the definition of a derivative, as written as $\lim\limits_{\Delta x \to 0} \frac{\Delta f}{\Delta x}$, it literally gives us the "change in f" for a "unit change in x". So that if we apply this same operation on itself, we would be finding the rate of change of the rate of change. If the rate of change of the rate of change slows down, we would get a hump shaped graph, while for the opposite where the rate increase, we would get a trough shaped graph. To put it in another way, *velocity* is the rate at which distance changes. However, the rate at which velocity changes is known as *acceleration*. Therefore if $y = f(t)$ denotes the position on a path at time t, $\dot{y} = \frac{dy}{dt}$ is velocity, so that acceleration is $\frac{d}{dt}\left(\frac{dy}{dt}\right)$.

Using the same notation as in our previous discussions using function $y = f(x)$, the *second derivative*, is thus the derivative of the derivative of $y = f(x)$, and is commonly denoted as,

$$\frac{dy^2}{dx^2},\ D^2y,\ y'',\ f'',\ D^2f,\ f^{(2)},\ f^{(2)}(x) \text{ or } \frac{df^2}{dx^2}$$

In turn, the *third derivative* is defined as the derivative of the second derivative of the function $y = f(x)$, $\frac{d}{dx}\left(\frac{d^2y}{dx^2}\right)$, and is commonly denoted as,

$$\frac{dy^3}{dx^3},\ D^3y,\ y''',\ f''',\ D^3f,\ f^{(3)},\ f^{(3)}(x) \text{ or } \frac{df^3}{dx^3}$$

This cascade of derivative can technically keep going, as long as the function is differentiable up to that level/order. Consider the following example, which is left for you to complete.

Example 8 *Compute $f^{(n)}(x)$ for $f(x) = x^4 + 7x^3 + 3x^2 - 9x + 1$, and where n is a positive integer.*

As hinted earlier, the higher derivatives can be used to tell you the shape of the function. To be clear, the definitions are as follows.

Definition 6 ***Concave Upward.*** *If a function f whose first derivative is increasing throughout the interval (a, b), we say the graph of the function is concave upward/convex. This thus mean that $f'' \geq 0$.*

In turn,

Definition 7 ***Concave Downward.*** *If a function f whose first derivative is decreasing throughout the interval (a, b), we say the graph of the function is concave downward/concave. This thus mean that $f'' \leq 0$.*

It is also possible that in a single function, you have portions of the graph exhibiting upward or downward concavity in turn. The point at which the graph switches is known as an *inflection point.*

Definition 8 ***Inflection point and inflection number.*** *Let f be a function and $c \in (a, b)$ such that $a < c < b$. For a continuous f on (a, b), where f is concave upward in (a, c) and concave downward in (c, b), or vice versa, then $(c, f(c))$ is an inflection point, and c is an inflection number.*

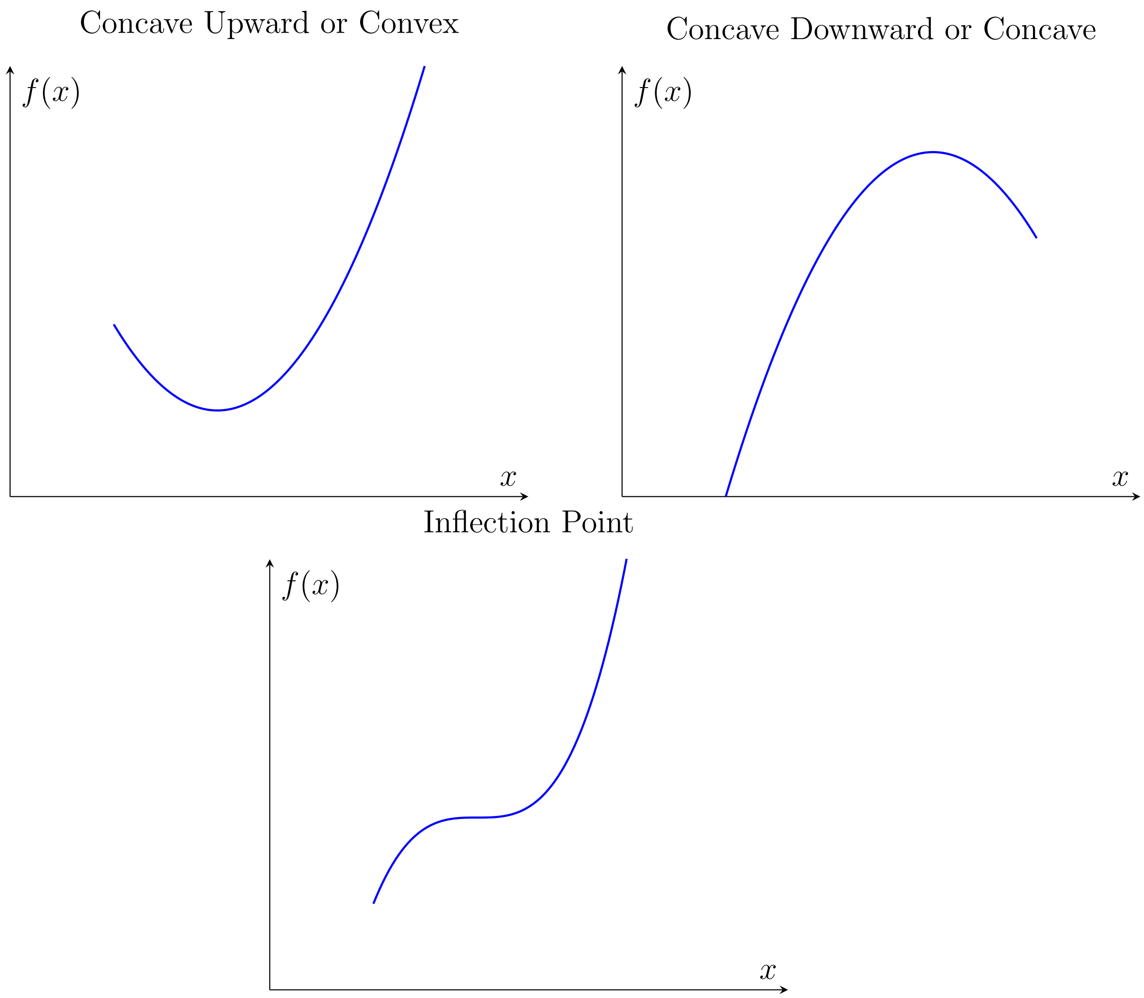
How do we pull all these definitions, and insights together in a useful manner? We have learned that within an open interval, under what circumstances, we could find a local extremum. Can derivatives be used to help us? We know based on Rolle's Theorem, that at an extremum, the first derivative must be zero, $f'(c) = 0$, for $c \in (a, b)$ where (a, b) is the domain of the function f. Looking at the diagrams in figure 5, it would further seem that the second derivative can be used to tell us the sort of extrema we have found.

Theorem 14 ***Second derivate test for relative maximum or minimum.*** *For a function f such that $f'(x)$ is defined on an open interval including c, assume that $f''(c)$ is defined. Then if $f'(c) = 0$, and $f''(c) < 0$, f achieves a local maximum at c. If instead $f'(c) = 0$, and $f''(c) > 0$, f achieves a local minimum at c.*

Proof. Suppose $f'(c) = 0$ and $f''(c) < 0$. Then for x sufficiently close to c,

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

Figure 5: Concavity



by the definition of a derivative (or more precisely a second derivative). Since $f'(c) = 0$, we have $\frac{f'(x)}{x-c} < 0$, so that if $x > c$ the inequality holds only if $f'(x) < 0$, while if $x < c$ the inequality holds only if $f'(x) > 0$. Thus as we move from left to right pass point c, the slope changes from positive to zero to negative, we have a local maximum at $x = c$.

The case for a local minimum can be similarly proved. ■

Let us bring what we have learned to bear with an economic example.

Example 9 *Let a firm's total revenue function be* $3q^{\frac{1}{2}} - 7q$*, and a total cost function of* $2q$*, where the choice of output is* q*. Find the optimal choice of* q*, and verify if this choice yield a local maximum.*

Solution 9 *Given the information the firm's profit function is,*

$$\Pi(q) = 3q^{\frac{1}{2}} - 7q - 2q$$

Since the optimum occurs where the slope/gradient of the function is zero, we can find the derivative of Π *with respect to* q*, and set it to zero.*

$$\begin{aligned}
\frac{d\Pi}{dq} &= \frac{3}{2}q^{-\frac{1}{2}} - 7 - 2 = 0 \\
\Rightarrow \frac{1}{\sqrt{q}} &= \frac{9 \times 2}{3} = 6 \\
\Rightarrow \sqrt{q} &= \frac{1}{6} \\
\Rightarrow q &= \frac{1}{36}
\end{aligned}$$

Further investigating whether this solution is a local maximum,

$$\frac{d\Pi^2}{d^2q} = -\frac{3}{4}q^{-\frac{3}{2}} < 0$$

So that the choice of $q = \frac{1}{36}$ *is indeed a local maximum, and the profit is maximized.*

Why is the maximum local and not global?

2.5.1 An Aside: The Binomial Theorem & Higher Order Derivatives

The *Binomial Theorem* asserts that,

$$(1+x)^n = 1 + nx + \frac{n(n-1)}{2!}x^2 + \cdots + \frac{n(n-1)\dots(n-k+1)}{k!}x^k + \cdots + \frac{n!}{n!}x^n$$

Where the operator ! is known as a *factorial*, and $k!$ is the product of the integers from 1 through k. Note that $0! = 1$, and the coefficient of x^k can be written as $\frac{n!}{k!(n-k)!}$ and is known as the *Binomial Coefficient*. It turns out that you can use higher order derivatives to show that the binomial coefficient is true.

Example 10 *Use higher derivatives to show that,*

$$(1+x)^4 = 1 + 4x + \frac{4.3}{2!}x^2 + \frac{4.3.2}{3!}x^3 + \frac{4.3.2}{4!}x^4$$

Solution 10 *The full expansion of* $(1+x)^4$ *would yield an equation of the form,*

$$a_0 + a_1x + a_2x^2 + a_3x^3 + a_4x^4 = (1+x)^4$$

If we take successive higher derivatives of the above equations, on both sides, we get,

$$\begin{aligned} a_1 + 2a_2 x + 3a_3 x^2 + 4a_4 x^3 &= 4(1+x)^3 \\ 2a_2 + 6a_3 x + 12a_4 x^2 &= 12(1+x)^2 \\ 6a_3 + 24a_4 x &= 24(1+x) \\ 24a_4 &= 24 \end{aligned}$$

From the above, the last equality gives us $a_4 = 1$*. Substituting this into the penultimate equality gives* $a_3 = 4$*. Substitute both into the second equality gives us* $a_2 = 6$*. Substitute all three into the first derivative gives us* $a_1 = 4$*, and we know* $a_0 = 1$*.*

You may use the same sequence of arguments to show that the Binomial Theorem holds for $(1+x)^n$.

2.6 Derivatives of Implicit Functions

Thus far we have dealt with functions of specific mathematical form. Consider a function given in the form of $y = f(x)$, say $y = f(x) = 3x^2$ is known as an explicit function, since we see the equational form, it is a quadratic equation. However, if instead we write the equation as $y - 3x^2 = 0$, it is now in implicitly form, since the relationship between y and x is only implicitly defined by $y - 3x^2 = 0$. We call such functions *implicit functions*. The last equation is often denoted generally by $F(x, y) = 0$. It should not be surmised that the right hand side of an implicit function must always be zero. For instance, $x^2 + 2y^2 = 25$ is likewise an implicit function, albeit of the form $F(x, y) = k$ where k is a constant. Of course we could solve it, say the last implicit function can be written in explicit forms, $y = \frac{\sqrt{25-x^2}}{2}$ or $y = -\frac{\sqrt{25-x^2}}{2}$. Nonetheless, it is possible to differentiate the implicit function, so as to discern their relationship without solving for $y = f(x)$ explicitly.

Example 11 *For* $x^2 + 2y^2 = 25$*, find* $\frac{dy}{dx}$*.*

Solution 11

$$\begin{aligned} \frac{d(x^2 + 2y^2)}{dx} &= \frac{d25}{dx} \\ \Rightarrow 2x + 4y\frac{dy}{dx} &= 0 \\ \Rightarrow \frac{dy}{dx} &= -\frac{x}{2y} \end{aligned}$$