INTRODUCTORY THEORY OF DIFFERENTIATION

This article is written for beginners. No previous knowledge of differentiation is assumed, and we shall treat the theory of limits and differentiation in a rigorous yet elementary way. If you have heard of the term "differentiation" and are mathematically curious enough to find out what it is, you should find it exciting to work through this article! (It may require some thought and hard work to understand everything, but yet if you're interested and patient enough you should be able to work through most of it.)

To begin with, in section 1 we introduce what differentiation is graphically by interpreting it as a process of finding slopes of graphs. Then in section 2, we give some examples to illustrate why we need differentiation. In section 3 we give a rigorous treatment of the theory of limits, which is essential for the theory of differentiation. Finally in section 4 we establish some rules of differentiation which you may find useful.

So our first question is: what is differentiation?

1. Graphical interpretation of Differentiation

Basically differentiation is a *process of finding slopes* of *tangents* to the graph of a given function. Given the graph of a function f (see figure 1), for a fixed x, if A, B represent the points $(x, f(x))$ and $(x+\Delta x, f(x+\Delta x))$ (here Δx must be read as one symbol in its own right just like any other symbols; Δ usually means change, so the symbol Δx suggests it represents a change in x), then the *slope* of AB is given by

$$\frac{f(x+\Delta x)-f(x)}{\Delta x}. \tag{1.1}$$

When Δx gets smaller, B gets closer to A, and when Δx is very small, we see that B nearly coincides with A. Then AB becomes nearly a *tangent* to the graph of f at A, and the quantity in (1.1) becomes a very good approximation to the slope of the *tangent* to the graph of f at A. "Taking limit" such that "Δx tends to zero", we get the slope of the *tangent* to the graph of f at A, given by

$$\lim_{\Delta x\to 0}\frac{f(x+\Delta x)-f(x)}{\Delta x}. \tag{1.2}$$

(The exact meaning of the quoted phrases and the sign $\lim_{\Delta x\to 0}$ in (1.2) will be explained in section 3.

For the present purpose, you may think of the sign $\lim_{\Delta x\to 0}$ to mean simply that "Δx is very close to zero".) Denote the quantity in (1.2) by *f'*(*x*) (because it depends on our initial choice of *x*), we get a new function *f'* which we call the **derivative** of the function *f*, and we see that for any *x*, *f'*(*x*) gives the slope of the tangent to the graph of *f* at *x*. Differentiation is then the process of finding the derivative of a given function.

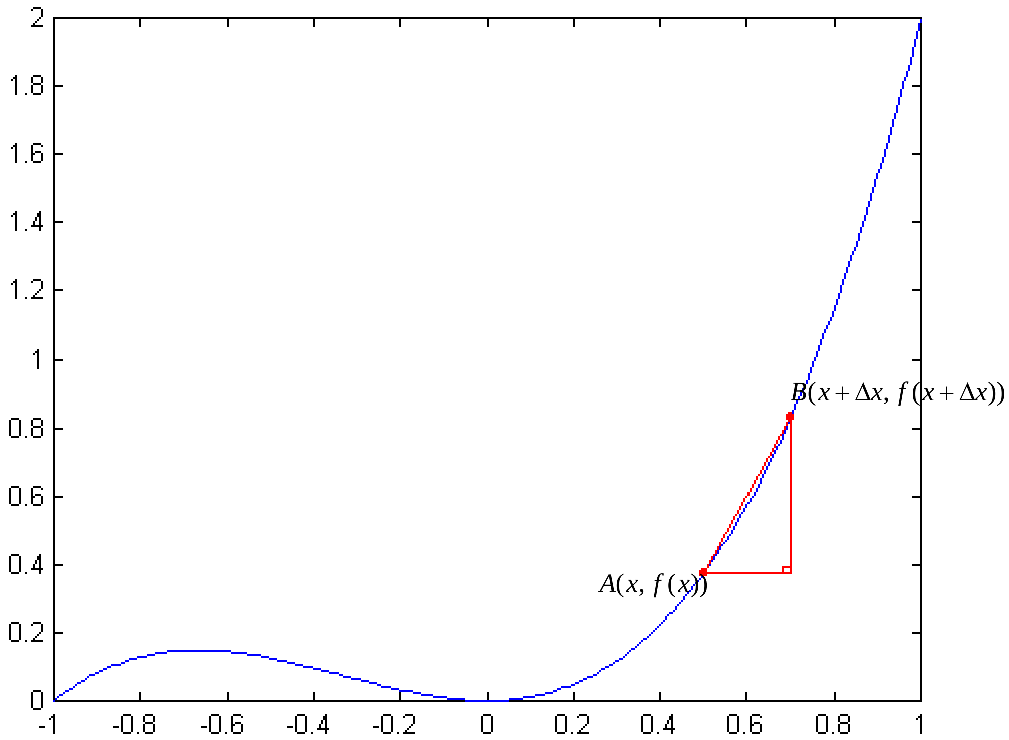

Figure 1

But then one may ask: why is differentiation important? Why do we need to find the derivative of a given function? Work through our next section, and you shall see a better reason why we need to study quantities like $\lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$ for a given *f* and *x* by giving examples of how such quantities naturally arise in nature.

2. Physical interpretation of Differentiation

Suppose a particle *P* moves along a straight line. Then given its displacement *s* (from a fixed reference point) as a function of time $t\,(t \geq 0)$, we know that the average velocity of *P* between time t_0 and $t_0 + \Delta t$ is given by "the change of displacement divided by the time taken", i.e.

$$\frac{s(t_0 + \Delta t) - s(t_0)}{\Delta t}$$

provided that both t_0 and $t_0 + \Delta t$ are non-negative (so that $s(t_0)$ and $s(t_0 + \Delta t)$ make sense).

When Δt is very small, or more precisely when "Δt tends to zero", we get the **instantaneous velocity** $v(t_0)$ of P at time t_0. We usually write

$$v(t_0) = \lim_{\Delta t \to 0} \frac{s(t_0 + \Delta t) - s(t_0)}{\Delta t}.$$

(Recall that the "$\lim_{\Delta t \to 0}$" sign denotes the process of "taking limit" as "Δt tends to zero".) You might have noticed how similar this formula is to (1.2). Do this for each $t \geq 0$, we get a function v such that $v(t)$ represents the instantaneous velocity of the particle P at time t, and

$$v(t) = \lim_{\Delta t \to 0} \frac{s(t + \Delta t) - s(t)}{\Delta t} \qquad (1.3)$$

for every $t \geq 0$. This function v is exactly the derivative of the function s.

Now suppose we want to quantify how fast air warms up when the sun rises. Let $t = 0$ denote the time of sunrise, and denote the air temperature at time $t\,(t \geq 0)$ by $T(t)$. Then the average rate of change in air temperature between time t_0 and $t_0 + \Delta t$ is given by

$$\frac{T(t_0 + \Delta t) - T(t_0)}{\Delta t}$$

provided that both t_0 and $t_0 + \Delta t$ are non-negative (so that $T(t_0)$ and $T(t_0 + \Delta t)$ make sense).

Again, when "Δt tends to zero", we get the **instantaneous rate of change of air temperature** $T'(t_0)$ at time t_0. We usually write

$$T'(t_0) = \lim_{\Delta t \to 0} \frac{T(t_0 + \Delta t) - T(t_0)}{\Delta t}.$$

Do this for each $t \geq 0$, we get a function T' such that $T'(t)$ represents the instantaneous rate of change of air temperature at time t, and

$$T'(t) = \lim_{\Delta t \to 0} \frac{T(t + \Delta t) - T(t)}{\Delta t} \qquad (1.4)$$

for every $t \geq 0$. Again, you may have realized that this function T' is exactly the derivative of the function T.

Note how similar formula (1.3) and (1.4) are. They are similar because both formula represents the *rate* at which a particular quantity (like displacement or air temperature) *changes* with time. Such **rate of change** problems arise naturally in many other areas of physics as well as sciences like economics, finance, biology, chemistry and so on. In treating them it is natural to come across expressions like

$$\lim_{\Delta t \to 0} \frac{f(t + \Delta t) - f(t)}{\Delta t}$$

where f is a function. Thus we see that differentiation is intimately related to rate of change problems which arise naturally in many aspects of everyday life. This explains why differentiation and the study of quantities like that in (1.2) is important.

Our next question is: what exactly do we mean by phrases like "taking limit" and "Δt tends to zero"?

3. Theory of Limits

Consider the function $f(x) = \frac{\sin x}{x}$. This function is undefined at $x = 0$, because division by zero is undefined. However, if we look at its graph (see figure 2 below), we see that as x gets very close to (but different from) 0, $f(x)$ gets very close to 1. Hence it seems that it makes sense to say "as x tends to 0, $f(x)$ tends to 1", or in our previous notation "$\lim_{x\to 0} f(x) = \lim_{x\to 0} \frac{\sin x}{x} = 1$".

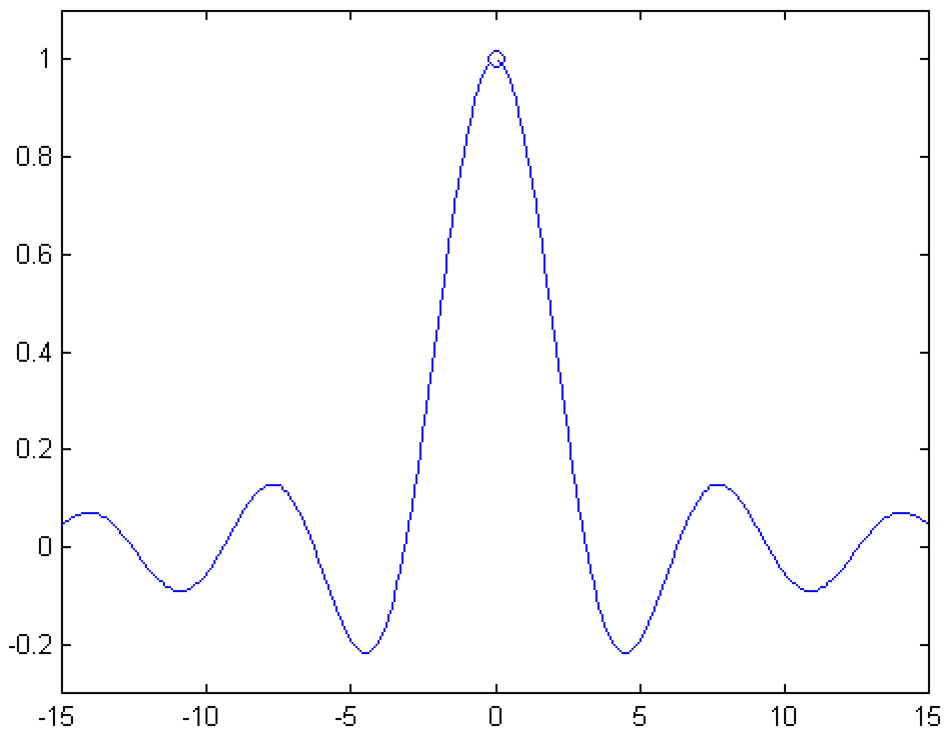

Figure 2: A graph of $f(x) = \frac{\sin x}{x}$

Consider another example: Look at the function $g(x) = \frac{2(x^2-4)}{x-2}$, we can never substitute x = 2 into the expression (because division by zero is not allowed). However, we see that for $x \neq 2$, $g(x)$ is simply $2x + 4$. Hence as x becomes very close to (but different from) 2, $g(x)$ becomes very close to $2\times 2+4=8$. So again it seems that it is reasonable to say that "as x tends to 2, $g(x)$ tends to 8", or in previous notation "$\lim_{x\to 2} g(x) = \lim_{x\to 2} \frac{2(x^2-4)}{x-2} = 8$".

But here comes a problem. What is written above is certainly not rigorous enough. A mathematician is never satisfied with such vague formulations. They will ask: you say "*f*(*x*) gets very close to 1" and "*g*(*x*) becomes very close to 8", so how close are they?

Let's look at the situation more closely. We said "*g*(*x*) is very close to 8", and we actually see that *g*(*x*) can be as close to 8 as we wish, as long as we take *x* close enough to (but different from) 2. For example, if you want *g*(*x*) to differ from 8 by a very small number say 0.1, you can simply take any *x* which is not equal to 2 and differs from 2 by less than say 0.05 (i.e. $x \neq 2$ and $|x-2| < 0.05$) and then for any such *x* *g*(*x*) will differ from the desired limit 8 by at most the prescribed 0.1 (check it!). If you want *g*(*x*) to differ from 8 by an even smaller number say 0.002, you can still do it by taking *x* even closer to (but not equal to) 2, say *x* which satisfies $x \neq 2$ and $|x-2| < 0.001$ and then such *x* will make *g*(*x*) differ from 8 by at most the prescribed 0.002. Actually if you want *g*(*x*) to differ from 8 by a very small positive number ε, no matter how small ε is, as long as it is positive, you can always take a small enough number $\delta > 0$ such that whenever $x \neq 2$ differ from 2 by less than that δ (this is a way of saying when you take *x* close enough to 2), the corresponding *g*(*x*) will then differ from 8 by at most ε (this is just saying *g*(*x*) is close enough to 8). Similarly, we see (from the graph of *f*) that if you want *f*(*x*) to differ from 1 by a very small positive number ε, no matter how small ε is, as long as it is strictly greater than zero, you can always find a small enough $\delta > 0$ such that whenever $x \neq 0$ differ from 0 by less than δ, the corresponding *f*(*x*) will then differ from 1 by at most ε. This motivates our definition of limits.

Definition 3.1.
Let *f* be a function. We say "*f*(*x*) tends to *L* as *x* tends to *a*", or equivalently "$\lim_{x\to a} f(x) = L$", if and only if the following holds:
For any $\varepsilon > 0$, there is a $\delta > 0$ such that whenever *x* satisfies $|x-a| < \delta$ and $x \neq a$, we have $|f(x) - L| < \varepsilon$.

This definition is not at all trivial; so don't worry if you feel that you don't understand the definition thoroughly. We illustrate the definition with several examples.

Example 3.1.

Let $f(x) = \frac{\sin x}{x}$ for $x \neq 0$ as above. Then by "$\lim_{x\to 0} f(x) = 1$" we mean that "given any positive number ε (the word "any" implicitly implies no matter how small ε is), we can always find a positive number δ such that whenever x satisfies $x \neq 0$ and $|x-0|=|x|<\delta$, then $|f(x)-1|<\varepsilon$". This seems to be true from the graph of *f*. Note that $|x-0|<\delta$ simply says that the distance between x and 0 is smaller than δ, and $|f(x)-1|<\varepsilon$ just says that the distance between $f(x)$ and 1 is smaller than ε.

Example 3.2.

Below we prove that if $g(x)=\dfrac{2(x^2-4)}{x-2}$ for $x \neq 2$, then $\lim_{x\to 2} g(x)=8$.

Proof. Note that $\lim_{x\to 2} g(x)=8$ means that given any positive number ε, we can always find a positive δ such that whenever $x \neq 2$ satisfies $|x-2|<\delta$, we have $|g(x)-8|<\varepsilon$. So let $\varepsilon>0$ be given, to complete the proof we need to find a positive δ (which may depend on ε) such that if x satisfies $x \neq 2$ and $|x-2|<\delta$, then $|g(x)-8|<\varepsilon$. Note that for $x \neq 2$ we have

$$|g(x)-8|=\left|\frac{2(x^2-4)}{x-2}-8\right|=|2(x+2)-8|=2|x-2|$$

so we see that when $\varepsilon>0$ is given, we can simply take $\delta=\dfrac{\varepsilon}{2}>0$ and then whenever x satisfies $x \neq 2$ and $|x-2|<\delta=\dfrac{\varepsilon}{2}$, we must have $|g(x)-8|=2|x-2|<2\cdot\dfrac{\varepsilon}{2}=\varepsilon$. This completes our proof.

Q.E.D.

Actually the key idea is here: to prove "$\lim_{x\to a} f(x)=L$", we need to estimate $|f(x)-L|$ and say we can make $|f(x)-L|$ very small, provided that x is very close to, but not equal to, a.

There is one more point to note: given an arbitrary function *f* and an arbitrary number *a*, it is NOT necessary that there is a number L such that $\lim_{x\to a} f(x)=L$. In that case we say $\lim_{x\to a} f(x)$ *does not exist*. For example, if *f* is the function such that

$$f(x)=\begin{cases}1 & \text{if } x>0\\ -1 & \text{if } x<0\end{cases},$$

then $\lim_{x\to 0} f(x)$ does not exist, because if there is a number L such that $\lim_{x\to 0} f(x)=L$, then by the definition of limit, for the specific $\varepsilon=0.5$, we must have a positive δ such that $|f(x)-L|<\varepsilon$ whenever $x \neq 0$ and $|x-0|<\delta$. By taking $x=\dfrac{\delta}{2}$, which clearly satisfies $x \neq 0$ and $|x-0|<\delta$, we see that L must satisfy $|L-1|=|f(x)-L|<0.5$, and similarly by taking $x=-\dfrac{\delta}{2}$, we see that L must satisfy $|L+1| = |f(x) - L| < 0.5$. No number L can satisfy both $|L-1| < 0.5$ and $|L+1| < 0.5$. This shows that no number L can satisfy $\lim_{x \to 0} f(x) = L$, or in other words $\lim_{x \to 0} f(x)$ does not exist. We can illustrate this graphically on the graph of f with the "jump" which occurs at $x = 0$ (see figure 3).

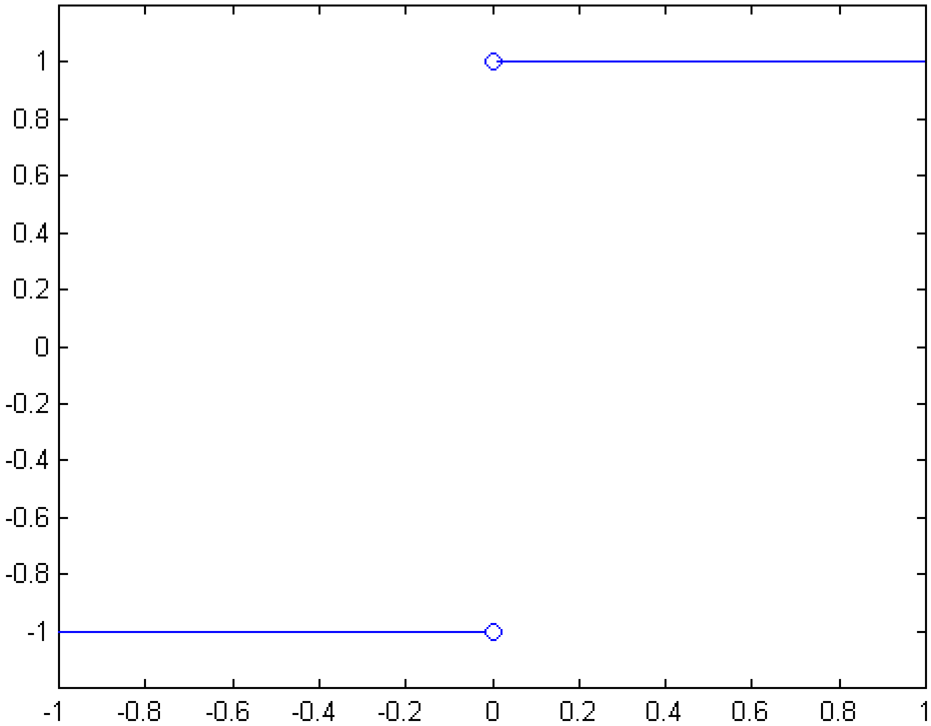

Figure 3: A graph of $f(x) = \begin{cases} 1 & \text{if } x > 0 \\ -1 & \text{if } x < 0 \end{cases}$

Below we prove some very useful properties of limits.

Proposition 3.1.
Let f, g be functions. If $f(x) = g(x)$ for all x except possibly at a, then either both $\lim_{x \to a} f(x)$ and $\lim_{x \to a} g(x)$ do not exist, or both of them exist and $\lim_{x \to a} f(x) = \lim_{x \to a} g(x)$.

In other words, when we talk about $\lim_{x \to a} f(x)$, we need not care about what $f(a)$ is. We do not even care about whether f is defined at a.

Proof. This is immediate from the definition of limit: Suppose $\lim_{x \to a} f(x)$ exists, let's call the limit L. Then let $\varepsilon > 0$ be given, by definition of limit, there must exist a positive number δ such that $|f(x) - L| < \varepsilon$ whenever x satisfies $x \neq a$ and $|x - a| < \delta$. Now for all such x, we have (since

$x \neq a$) that $f(x) = g(x)$, so whenever x satisfies $x \neq a$ and $|x-a| < \delta$, we must have $|g(x)-L| = |f(x)-L| < \varepsilon$. This shows that $\lim_{x \to a} g(x)$ exists and is equal to L.

Q.E.D.

Theorem 3.2.

Let f, g be functions. If $\lim_{x \to a} f(x)$ and $\lim_{x \to a} g(x)$ both exists, then

1. $\lim_{x \to a}(f+g)(x)$ exists and $\lim_{x \to a}(f+g)(x) = \lim_{x \to a} f(x) + \lim_{x \to a} g(x)$;
2. $\lim_{x \to a}(f \cdot g)(x)$ exists and $\lim_{x \to a}(f \cdot g)(x) = \lim_{x \to a} f(x) \cdot \lim_{x \to a} g(x)$;
3. If $\lim_{x \to a} g(x) \neq 0$ then $\lim_{x \to a} \frac{f}{g}(x)$ exists and $\lim_{x \to a} \frac{f}{g}(x) = \frac{\lim_{x \to a} f(x)}{\lim_{x \to a} g(x)}$.

Proof. Suppose $\lim_{x \to a} f(x)$ and $\lim_{x \to a} g(x)$ both exists, let $L_1 = \lim_{x \to a} f(x)$ and $L_2 = \lim_{x \to a} g(x)$.

1. The key idea is to make use of the *triangle inequality*:

   $|a+b| \leq |a| + |b|$ for all real numbers a, b. (Prove it yourself!)

   Let $\varepsilon > 0$ be given. Then $\varepsilon / 2 > 0$ too, so by definition of limit we have

   I. there exists a $\delta_1 > 0$ such that whenever x satisfies $x \neq a$ and $|x-a| < \delta_1$, we have $|f(x) - L_1| < \varepsilon / 2$; and

   II. there exists a $\delta_2 > 0$ such that whenever x satisfies $x \neq a$ and $|x-a| < \delta_2$, we have $|g(x) - L_2| < \varepsilon / 2$.

   Now let $\delta = \min\{\delta_1, \delta_2\}$, then we see that $\delta > 0$ because both δ_1, δ_2 are positive. Furthermore, whenever x satisfies $x \neq a$ and $|x-a| < \delta$, we have

   III. $|f(x) - L_1| < \varepsilon / 2$ by I (because $\delta \leq \delta_1$ by definition of δ); and

   IV. $|g(x) - L_2| < \varepsilon / 2$ by II (because $\delta \leq \delta_2$ by definition of δ).

   Combining III and IV we have

   $$|(f+g)(x) - (L_1 + L_2)| = |(f(x) - L_1) + (g(x) - L_2)| \leq |f(x) - L_1| + |g(x) - L_2| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \quad (3.1)$$

   where the first inequality follows from triangle inequality and the second inequality follows from III and IV. Since (3.1) is true for all x which satisfies $x \neq a$ and $|x-a| < \delta$, we have

   $\lim_{x \to a}(f+g)(x)$ exists and $\lim_{x \to a}(f+g)(x) = L_1 + L_2 = \lim_{x \to a} f(x) + \lim_{x \to a} g(x)$.

2. We want to prove $\lim_{x \to a}(f \cdot g)(x) = L_1 L_2$, so we try to estimate $|(f \cdot g)(x) - L_1 L_2|$ for x close to, but not equal to a: note that

$$\begin{aligned}
&|(f\cdot g)(x)-L_1L_2|\\
&=|f(x)g(x)-L_1L_2|\\
&=|(f(x)g(x)-f(x)L_2)+(f(x)L_2-L_1L_2)|\\
&\le|f(x)g(x)-f(x)L_2|+|f(x)L_2-L_1L_2|\\
&=|f(x)||g(x)-L_2|+|L_2||f(x)-L_1|
\end{aligned} \tag{3.2}$$

for all *x*, where $|f(x)-L_1|$, $|g(x)-L_2|$ are quantities which can be made very small by choosing only *x* very close to, but not equal to *a*. Now $|L_2|$ is a fixed number, so the second term in the last line of (3.2) can be made very small, as long as *x* is very close to (but not equal to) *a*. So if we can show that $|f(x)|$ is not too large for *x* close to (but not equal to) *a*, then the first term would also be small for such *x* and we are almost done. Hence we are motivated to prove the following first:

There exists a positive number δ_0 and a positive number *M* such that $|f(x)|<M$ for all *x* satisfying $x\neq a$ and $|x-a|<\delta_0$. --(*)

Proof of (*): Let $\varepsilon_0=1$, then $\varepsilon_0>0$, so by definition of limit we see that there exists $\delta_0>0$ such that $|f(x)-L_1|<\varepsilon_0=1$ whenever *x* satisfies $x\neq a$ and $|x-a|<\delta_0$. Now for all such *x* we have

$$|f(x)|\le|f(x)-L_1|+|L_1|$$

(triangle inequality again!). Hence $|f(x)|<1+|L_1|$ whenever *x* satisfies $x\neq a$ and $|x-a|<\delta_0$. Take $M=1+|L_1|$, we see that δ_0, *M* are both positive and that $|f(x)|<M$ and for all *x* satisfying $x\neq a$ and $|x-a|<\delta_0$, and hence (*) is proved.

We go on to complete our proof:

We want to show $\lim_{x\to a}(f\cdot g)(x)$ exists and $\lim_{x\to a}(f\cdot g)(x)=L_1L_2$, so let $\varepsilon>0$ be given. Then $\varepsilon/2M$ is also positive, where *M* is as given by (*), so by definition of limit, we have

I. there is a positive number δ_2 such that $|g(x)-L_2|<\varepsilon/2M$ whenever *x* satisfies $x\neq a$ and $|x-a|<\delta_2$, and choose smaller δ_2 if necessary we may assume $\delta_2<\delta_0$.

Furthermore, $\dfrac{\varepsilon}{2(1+|L_2|)}$ is also positive, so by definition of limit again, we have

II. there is a positive number δ_1 such that $|f(x)-L_1|<\dfrac{\varepsilon}{2(1+|L_2|)}$ whenever *x* satisfies $x\neq a$ and $|x-a|<\delta_1$, and similar to I we may assume $\delta_1<\delta_0$.

Now let $\delta=\min\{\delta_1,\delta_2\}$, then we see that $\delta>0$ because both δ_1,δ_2 are. Furthermore, whenever *x* satisfies $x\neq a$ and $|x-a|<\delta$, we have

III. $|g(x)-L_2|<\varepsilon/2M$ by I (because $\delta\le\delta_1$ by definition of δ) ;

IV. $|f(x)-L_1|<\dfrac{\varepsilon}{2(1+|L_2|)}$ by II (because $\delta\le\delta_2$ by definition of δ) ; and

V. $|f(x)|<M$ by (*) (because both $\delta_1,\delta_2<\delta_0$ implies $\delta<\delta_0$).

Combining III, IV and V, we have

$$\begin{aligned} &|(f\cdot g)(x)-L_1L_2| \\ &\le |f(x)||g(x)-L_2|+|L_2||f(x)-L_1| \\ &< M\cdot\frac{\varepsilon}{2M}+|L_2|\frac{\varepsilon}{2(1+|L_2|)} \\ &<\varepsilon \end{aligned} \tag{3.3}$$

where the first inequality follows from (3.2) and the second inequality follows from III, IV and V. Since (3.3) is true for all x satisfying $x\neq a$ and $|x-a|<\delta$, we have $\lim_{x\to a}(f\cdot g)(x)$ exists and $\lim_{x\to a}(f\cdot g)(x)=L_1L_2$. This completes our proof.

3. Again we try to make a rough estimation first: note $L_2\neq 0$ is given, so if $g(x)\neq 0$ then by triangle inequality again we have

$$\begin{aligned} &\left|\frac{f}{g}(x)-\frac{L_1}{L_2}\right|=\left|\frac{f(x)}{g(x)}-\frac{L_1}{L_2}\right|=\frac{|f(x)L_2-L_1g(x)|}{|g(x)||L_2|} \\ &=\frac{|f(x)L_2-L_1L_2+L_1L_2-L_1g(x)|}{|g(x)||L_2|} \\ &\le\frac{|f(x)L_2-L_1L_2|+|L_1L_2-L_1g(x)|}{|g(x)||L_2|} \\ &=\frac{|L_2||f(x)-L_1|+|L_1||g(x)-L_2|}{|g(x)||L_2|} \end{aligned} \tag{3.4}$$

where $|f(x)-L_1|$, $|g(x)-L_2|$ are quantities which can be made very small by choosing only x very close to, but not equal to a. Now $|L_1|$, $|L_2|$ are fixed numbers, so the numerator in the last line of (3.4) can be made very small, as long as x is very close to (but not equal to) a. So if we can show that $|g(x)|$ is not too small for x close to (but not equal to) a, then the whole fraction would also be small for such x and we are almost done. Hence we are motivated to prove the following first:

There exists a positive number δ_0 and a positive number M such that $|g(x)|>M$ for all x satisfying $x\neq a$ and $|x-a|<\delta_0$. --(*)

Proof of (*): Note $L_2\neq 0$ is given, so $|L_2|>0$. Let $\varepsilon_0=|L_2/2|$, then $\varepsilon_0>0$, so by definition of limit we see that there exists $\delta_0>0$ such that $|g(x)-L_2|<\varepsilon_0=|L_2/2|$ whenever x satisfies $x\neq a$ and $|x-a|<\delta_0$. This shows for all such x we have

$$|g(x)|>|L_2/2|.$$

(why?) Take $M=|L_2/2|$, we see that δ_0, M are both positive and that $|g(x)|>M$ and for all x satisfying $x\neq a$ and $|x-a|<\delta_0$, and hence (*) is proved.

Our proof is completed below:

We want to show $\lim_{x\to a}\frac{f}{g}(x)$ exists and $\lim_{x\to a}\frac{f}{g}(x)=\frac{L_1}{L_2}$, so let $\varepsilon>0$ be given. Then $M\varepsilon/2$ is also positive, where M is as given by (*), so by definition of limit, we have

I. there is a positive number δ_1 such that $|f(x)-L_1|<M\varepsilon/2$ whenever x satisfies $x\neq a$ and $|x-a|<\delta_1$, where we assume $\delta_1<\delta_0$.

Furthermore, $\dfrac{M|L_2|\varepsilon}{2(1+|L_1|)}$ is also positive, so by definition of limit again, we have

II. there is a positive number δ_2 such that $|g(x)-L_2|<\dfrac{M|L_2|\varepsilon}{2(1+|L_1|)}$ whenever x satisfies $x\neq a$ and $|x-a|<\delta_2$, where we assume $\delta_2<\delta_0$.

Now let $\delta=\min\{\delta_1,\delta_2\}$, then we see that $\delta>0$ because both δ_1,δ_2 are positive. Furthermore, whenever x satisfies $x\neq a$ and $|x-a|<\delta$, we have

III. $|f(x)-L_1|<M\varepsilon/2$ by I (because $\delta\le\delta_1$ by definition of δ) ;

IV. $|g(x)-L_2|<\dfrac{M|L_2|\varepsilon}{2(1+|L_1|)}$ by II (because $\delta\le\delta_2$ by definition of δ) ; and

V. $|g(x)|>M$ by (*) (because both $\delta_1,\delta_2<\delta_0$ implies $\delta<\delta_0$).

Combining III, IV and V, we have

$$\begin{aligned}&\left|\frac{f}{g}(x)-\frac{L_1}{L_2}\right|\\ &\le\frac{|L_2||f(x)-L_1|+|L_1||g(x)-L_2|}{|g(x)||L_2|}\\ &<\frac{M\varepsilon}{2}\cdot\frac{1}{M}+|L_1|\frac{M\varepsilon}{2(1+|L_1|)}\cdot\frac{1}{M}\\ &<\varepsilon\end{aligned}\tag{3.5}$$

where the first inequality follows from (3.4) and the second inequality follows from III, IV and V. Since (3.5) is true for all x satisfying $x\neq a$ and $|x-a|<\delta$, we have $\lim_{x\to a}\frac{f}{g}(x)$ exists and $\lim_{x\to a}\frac{f}{g}(x)=\frac{L_1}{L_2}$. Our proof is hence complete.

Q.E.D.

Proposition 3.3.

For all positive integers n, $\lim_{x\to a}x^n$ exists and $\lim_{x\to a}x^n=a^n$.

Proof. Note that $\lim_{x\to a}x$ exists and is equal to a, because given any $\varepsilon>0$, we can simply take $\delta=\varepsilon$, and then we will have $\delta>0$ as well as $|x-a|<\varepsilon$ whenever x satisfies $|x-a|<\delta$. Now for any positive integers n, since $x^n=\underbrace{x\cdot x\cdot\cdots\cdot x\cdot x}_{n\text{ terms}}$, so by repeated use of Theorem 3.2.2 we have $\lim_{x\to a}x^n$ exists and $\lim_{x\to a}x^n=\underbrace{(\lim_{x\to a}x)\cdot(\lim_{x\to a}x)\cdot\cdots\cdot(\lim_{x\to a}x)\cdot(\lim_{x\to a}x)}_{n\text{ terms}}=\underbrace{a\cdot a\cdot\cdots\cdot a\cdot a}_{n\text{ terms}}=a^n$. This completes our proof.

(Can you give a more rigorous presentation of this proof using mathematical induction?)

Q.E.D.

Exercise

You are highly advised to go through the following exercises to check that you have firmly grasp the right concepts before you move on to the next section where we shall use the above concepts frequently in establishing the rigorous way to differentiation. You should always state clearly where you have used the propositions and theorem mentioned above.

3.1 Explain why the following limits exist and evaluate them:

(a) $\lim_{x\to 3}(x^2+3x+1)$

(b) $\lim_{x\to -2}(x+1)(2x^2-x^3)$

(c) $\lim_{x\to -1}(x-3)(2x^5+3x^3)(x^2+1)$

(d) $\lim_{x\to 0} f(x)$, where $f(x)=\dfrac{x^2+3x^3}{x^2}$ for $x\neq 0$

(e) $\lim_{x\to -2}\dfrac{x^2-4}{x+2}$

(f) $\lim_{x\to 2}\dfrac{x^3-8}{(x-2)(x+1)}$

3.2 Prove that for all real numbers *c*, $\lim_{x\to a} c$ exists and $\lim_{x\to a} c = c$. (Note that here $\lim_{x\to a} c$ means the limit of a function $g(x)$ as *x* tends to *a*, where *g* is a function with $g(x)=c$ for all *x*.)

3.3 Prove that for all positive integers *n* and for all $a\neq 0$, $\lim_{x\to a}\dfrac{1}{x^n}$ exists and $\lim_{x\to a}\dfrac{1}{x^n}=\dfrac{1}{a^n}$.

3.4 Prove that for all real numbers *a*, $\lim_{\Delta x\to 0}(a+\Delta x)^k$ exists and $\lim_{\Delta x\to 0}(a+\Delta x)^k=a^k$ for all positive integers *k*.

3.5 Prove that if *g* is a function and *x* is such that $\lim_{\Delta x\to 0}\dfrac{g(x+\Delta x)-g(x)}{\Delta x}$ exists, then $\lim_{\Delta x\to 0} g(x+\Delta x)$ exists and $\lim_{\Delta x\to 0} g(x+\Delta x)=g(x)$.

3.6 Prove that if *g* is a function and *x* is such that $\lim_{\Delta x\to 0} g(x+\Delta x)$ exists and $\lim_{\Delta x\to 0} g(x+\Delta x)\neq 0$, then there is a positive number δ such that $g(y)\neq 0$ for all *y* satisfying $y\neq x$ and $|y-x|<\delta$.

4. Theory of Differentiation

As mentioned in section 1, differentiation is the process of finding the derivatives of a given function, and the derivative of a function *f* at *x* is given by

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

Actually this is not quite true; for a general function *f* and an arbitrary *x* there is no guarantee that the above limit exists. The following definition gives a more accurate picture of what is happening here:

Definition 4.1.

Let *f* be a function. We say *f* is **differentiable at *x*** if the limit

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

exists. In that case we denote the limit by $f'(x)$ and say $f'(x)$ is the **derivative of *f* at *x***. Furthermore, if *f* is differentiable at every *x* where *f* is defined, then we say simply that *f* is **differentiable** and then f' can then be seen as a new function defined at every *x* where *f* is defined, called the **derivative** of *f*.

Note that geometrically $f'(a)$ gives the "*slope*" of the graph of *f* at *a*. (See section 1.)

We give an example which involves the differentiation of some simple polynomials.

Proposition 4.1.

1. Let *f* be a constant function. Then *f* is differentiable, and $f'(x) = 0$ for all *x*.
2. Let *n* be a positive integer and let $f(x) = x^n$ for all *x*. Then *f* is differentiable, and $f'(x) = nx^{n-1}$ for all *x*.

Geometrically the second assertion means that this means the slope of the graph of $y = x^n$ at $x = a$ is given by na^{n-1}. For example, the slope of the graph of $y = x^2$ at $x = 3$ is $2 \times 3^{2-1} = 6$.

Proof.

1. Suppose *f* is a constant function. Then there is a constant *c* such that $f(x) = c$ for all *x*. Now for any fixed *x*, we have $\frac{f(x+\Delta x) - f(x)}{\Delta x} = \frac{c-c}{\Delta x} = 0$ for all $\Delta x \neq 0$. By Exercise 3.2, we see that $\lim_{x \to 0} 0$ exists and is equal to 0. So by Proposition 3.1, we see that $\lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$

exists and $\lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} = \lim_{\Delta x \to 0} 0 = 0$. This means *f* is differentiable at *x* and

$f\,'(x) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} = 0$. Since the above is true for all *x*, we have *f* is differentiable,

and $f\,'(x) = 0$ for all *x*.

2. Let *n* be a positive integer and let $f(x) = x^n$ for all *x*. For any fixed *a*, we have

$\frac{f(a+\Delta x) - f(a)}{\Delta x} = \frac{(a+\Delta x)^n - a^n}{\Delta x} = (a+\Delta x)^{n-1} + (a+\Delta x)^{n-2} a + \ldots + (a+\Delta x) a^{n-2} + a^{n-1}$ for all

$\Delta x \neq 0$ (why?). Now by Exercise 3.4, $\lim_{\Delta x \to 0} (a+\Delta x)^k$ exists and $\lim_{\Delta x \to 0} (a+\Delta x)^k = a^k$ for all

positive integers *k*. So by repeated use of Theorem 3.2.1 and 3.2.2 we have

$\lim_{\Delta x \to 0} [(a+\Delta x)^{n-1} + (a+\Delta x)^{n-2} a + \ldots + (a+\Delta x) a^{n-2} + a^{n-1}]$ exists and

$\lim_{\Delta x \to 0} [(a+\Delta x)^{n-1} + (a+\Delta x)^{n-2} a + \ldots + (a+\Delta x) a^{n-2} + a^{n-1}] = \underbrace{a^{n-1} + a^{n-1} + \ldots + a^{n-1}}_{n \text{ terms}} = na^{n-1}$. Hence by

Proposition 3.1 we see that $\lim_{\Delta x \to 0} \frac{f(a+\Delta x) - f(a)}{\Delta x}$ exists and $\lim_{\Delta x \to 0} \frac{f(a+\Delta x) - f(a)}{\Delta x} = na^{n-1}$.

This shows *f* is differentiable at *a* and $f\,'(a) = \lim_{\Delta x \to 0} \frac{f(a+\Delta x) - f(a)}{\Delta x} = na^{n-1}$. Since this is true

for all *a*, we have *f* is differentiable, and $f\,'(x) = nx^{n-1}$ for all *x* and we are done.

Q.E.D.

We have a useful theorem concerning the derivatives of sums, products and quotients of differentiable functions (c.f. Theorem 3.2).

Proposition 4.2.

Let *f*, *g* be functions. If *f* and *g* are both differentiable at *x*, then

1. $f + g$ is differentiable at *x* and $(f+g)'(x) = f\,'(x) + g\,'(x)$;
2. $f \cdot g$ is differentiable at *x* and $(f \cdot g)'(x) = f\,'(x) g(x) + f(x) g\,'(x)$;
3. If $g(x) \neq 0$ then $\frac{f}{g}$ is differentiable at *x* and $(\frac{f}{g})'(x) = \frac{f\,'(x) g(x) - f(x) g\,'(x)}{[g(x)]^2}$.

Proof. Suppose both *f* and *g* are differentiable at *x*. Then by definition and Exercise 3.5 we have

I. $\lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}$ exists and $\lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} = f\,'(x)$;

II. $\lim_{\Delta x \to 0} \frac{g(x+\Delta x) - g(x)}{\Delta x}$ exists and $\lim_{\Delta x \to 0} \frac{g(x+\Delta x) - g(x)}{\Delta x} = g\,'(x)$; and

III. $\lim_{\Delta x \to 0} g(x+\Delta x)$ exists and $\lim_{\Delta x \to 0} g(x+\Delta x) = g(x)$.

1. Note that for $\Delta x \neq 0$, we have

$$\frac{(f+g)(x+\Delta x)-(f+g)(x)}{\Delta x}=\frac{f(x+\Delta x)-f(x)}{\Delta x}+\frac{g(x+\Delta x)-g(x)}{\Delta x} \tag{4.1}$$

Now by Theorem 3.2.1, from I and II we see that $\lim_{\Delta x\to 0}[\frac{f(x+\Delta x)-f(x)}{\Delta x}+\frac{g(x+\Delta x)-g(x)}{\Delta x}]$ exists and is equal to $f'(x)+g'(x)$. Hence by (4.1) and Proposition 3.1 we have $\lim_{\Delta x\to 0}\frac{(f+g)(x+\Delta x)-(f+g)(x)}{\Delta x}$ exists and $\lim_{\Delta x\to 0}\frac{(f+g)(x+\Delta x)-(f+g)(x)}{\Delta x}=f'(x)+g'(x)$.

This says $f+g$ is differentiable at x and $(f+g)'(x)=f'(x)+g'(x)$.

2. Note that for $\Delta x \neq 0$, we have

$$\begin{aligned}&\frac{(f\cdot g)(x+\Delta x)-(f\cdot g)(x)}{\Delta x}\\&=\frac{f(x+\Delta x)g(x+\Delta x)-f(x)g(x)}{\Delta x}\\&=\frac{f(x+\Delta x)-f(x)}{\Delta x}g(x+\Delta x)+f(x)\frac{g(x+\Delta x)-g(x)}{\Delta x}\end{aligned} \tag{4.2}$$

(check it!) Now I, II and III together with Theorem 3.2.1 and 3.2.2 imply that $\lim_{\Delta x\to 0}[\frac{f(x+\Delta x)-f(x)}{\Delta x}g(x+\Delta x)+f(x)\frac{g(x+\Delta x)-g(x)}{\Delta x}]$ exists and is equal to $f'(x)g(x)+f(x)g'(x)$. Proposition 3.1 and (4.2) then shows that $\lim_{\Delta x\to 0}\frac{(f\cdot g)(x+\Delta x)-(f\cdot g)(x)}{\Delta x}$ exists and $\lim_{\Delta x\to 0}\frac{(f\cdot g)(x+\Delta x)-(f\cdot g)(x)}{\Delta x}=f'(x)g(x)+f(x)g'(x)$. This says $f\cdot g$ is differentiable at x and $(f\cdot g)'(x)=f'(x)g(x)+f(x)g'(x)$.

3. Observe that by II, $\lim_{\Delta x\to 0}\frac{g(x+\Delta x)-g(x)}{\Delta x}$ exists and so by Exercise 3.5, $\lim_{\Delta x\to 0}g(x+\Delta x)$ exists and $\lim_{\Delta x\to 0}g(x+\Delta x)=g(x)$. Now $g(x)\neq 0$, so $\lim_{\Delta x\to 0}g(x+\Delta x)\neq 0$, hence by Exercise 3.6 there is a positive number δ such that $g(y)\neq 0$ for all y satisfying $y\neq x$ and $|y-x|<\delta$.
Therefore for Δx satisfying $\Delta x\neq 0$ and $|\Delta x|<\delta$, we have $g(x+\Delta x)\neq 0$ and hence

$$\begin{aligned}&\frac{(\frac{f}{g})(x+\Delta x)-(\frac{f}{g})(x)}{\Delta x}\\&=\frac{\frac{f(x+\Delta x)}{g(x+\Delta x)}-\frac{f(x)}{g(x)}}{\Delta x}\\&=\frac{1}{g(x)g(x+\Delta x)}\frac{f(x+\Delta x)g(x)-f(x)g(x+\Delta x)}{\Delta x}\\&=\frac{1}{g(x)g(x+\Delta x)}[\frac{f(x+\Delta x)-f(x)}{\Delta x}g(x)-f(x)\frac{g(x+\Delta x)-g(x)}{\Delta x}]\end{aligned} \tag{4.3}$$

(check it!) Note that here $\lim_{\Delta x\to 0} g(x+\Delta x)\neq 0$, so I, II and III together with Theorem 3.2.1, 3.2.2 and 3.2.3 imply that $\lim_{\Delta x\to 0}\frac{1}{g(x)g(x+\Delta x)}[\frac{f(x+\Delta x)-f(x)}{\Delta x}g(x)-f(x)\frac{g(x+\Delta x)-g(x)}{\Delta x}]$ exists and is equal to $\frac{f'(x)g(x)-f(x)g'(x)}{[g(x)]^2}$. Proposition 3.1 and (4.3) then implies that

$\lim_{\Delta x\to 0}\frac{(\frac{f}{g})(x+\Delta x)-(\frac{f}{g})(x)}{\Delta x}$ exists and $\lim_{\Delta x\to 0}\frac{(\frac{f}{g})(x+\Delta x)-(\frac{f}{g})(x)}{\Delta x}=\frac{f'(x)g(x)-f(x)g'(x)}{[g(x)]^2}$.

This says $\frac{f}{g}$ is differentiable at *x* and $(\frac{f}{g})'(x)=\frac{f'(x)g(x)-f(x)g'(x)}{[g(x)]^2}$.

Q.E.D.

The following Corollary follows from (b) of the above proposition.

Corollary 4.3.
Let *f* be a function and *c* be a constant. If *f* is differentiable at *x*, then cf is also differentiable at *x* and $(cf)'(x)=cf'(x)$.

Proof. In Proposition 4.2.2, let *g* be defined by $g(y)=c$ for all *y*. Then by Proposition 4.1.1, *g* is differentiable at *x* and $g'(x)=0$. The result then follows from Proposition 4.2.2.

Q.E.D.

Example 4.1. If a function *g* is defined by $g(x)=-7x^2$, then *g* is differentiable at *x* for all real *x* and $g'(x)=-7(2x^{2-1})=-14x$ by Proposition 4.1.2 and Corollary 4.3. So if a function *f* is defined by $f(x)=x^5-7x^2$, then *f*, being the sum of two differentiable functions, is differentiable at any real *x* and $f'(x)=5x^4-14x$ by Proposition 4.1.2 and 4.2.1. Hence for example the slope of *f* at (1,-2) is $f'(1)=5\times 1^4-14\times 1=-9$. Indeed we can differentiate any polynomial now.

Example 4.2. If a function *f* is defined by $f(x)=\frac{1}{x}$ for non-zero real-values of *x*, then for all such *x*, we have *f* is differentiable at *x* and that $f'(x)=\frac{x\cdot 0-1\cdot 1}{x^2}=-\frac{1}{x^2}$ by Propositions 4.2.3 and 4.1.2. For example, we see that $f'(2)=-\frac{1}{2^2}=-\frac{1}{4}$, which says the slope of *f* at (2, $\frac{1}{2}$) is $-\frac{1}{4}$.

To conclude this article let us mention the intuitive idea of how differentiation can be used to find the maximum and minimum of a function. Look at the figure 4, we see that *f* attains its maximum and minimum at x_1 and x_2 respectively, where it seems $f'(x_1)=f'(x_2)=0$. A little bit thinking might convince you that this seems reasonable, because the slope of *f* usually vanishes where *f* attains a maximum or a minimum. But must it be that $f'(t)=0$ if *f* attains its maximum or minimum at t? The answer is no, because the function can be of the form in figure 5! What's wrong there?

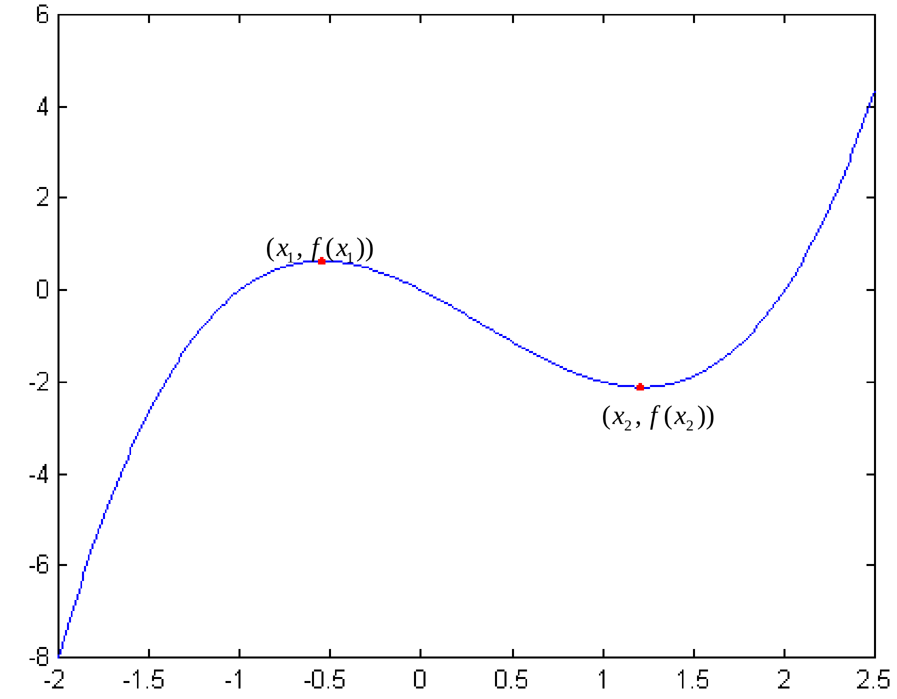

Figure 4

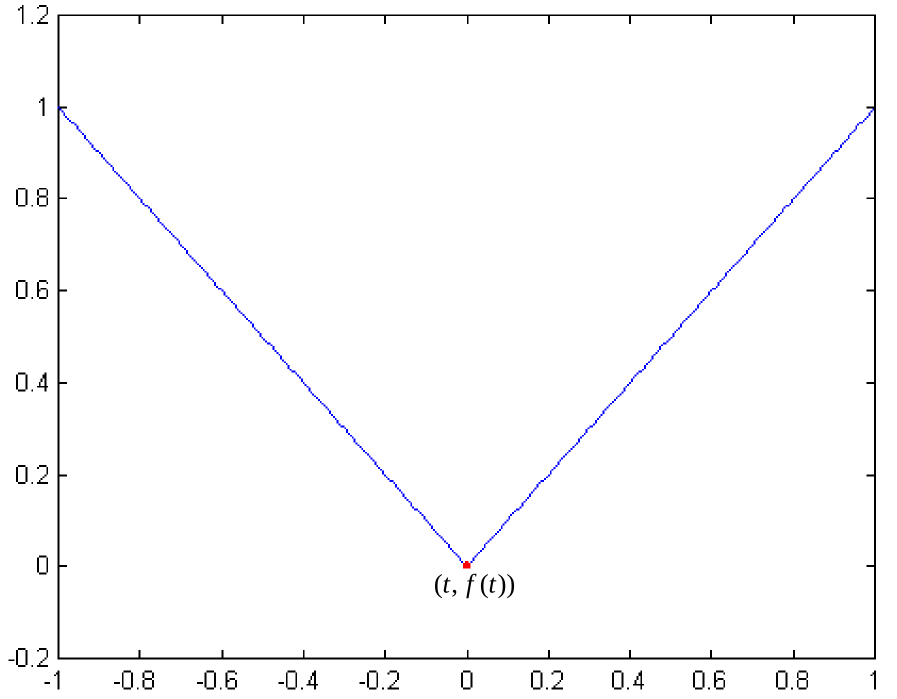

Figure 5

The problem is that in figure 5 $f'(t)$ does not exist (see figure). In other words, *f* is not differentiable at *t*. But the following is true for functions defined on the whole real line: For a function *f* defined on the whole real line, if *f* attains its maximum or minimum at *t* and *f* is differentiable near *t*, then $f'(t) = 0$. Actually the words "maximum" and "minimum" here can be replaced by "local maximum" and "local minimum" respectively, because all we are concerned with are local behaviour of *f* near the point *t*. Hence to find the local maximums or minimums of the function $f(x) = x^4 - 2x^2$, since it is differentiable on the whole real line, we first find *t* where $f'(t) = 0$. This is easy: it is easily verified that $f'(x) = 4x^3 - 4x$ for all *x* so $f'(t) = 0$ implies $4t^3 - 4t = 0$ and hence *t* = -1, 0 or 1. Hence the local maximums or minimums of this *f* can only occur at *x* = -1, 0 or 1. (And from a graph of it we can easily tell which is local maximum and which are local minimums.)

Conversely, suppose $f'(t) = 0$, does it necessarily imply that the function *f* attains a local maximum or local minimum at *t*? The answer is again no, as the following example reveal: The function $f(x) = x^3$ is clearly differentiable and $f'(0) = 3 \times 0^2 = 0$ by Proposition 4.1.2, but 0 is neither a local maximum nor a local minimum of *f*.

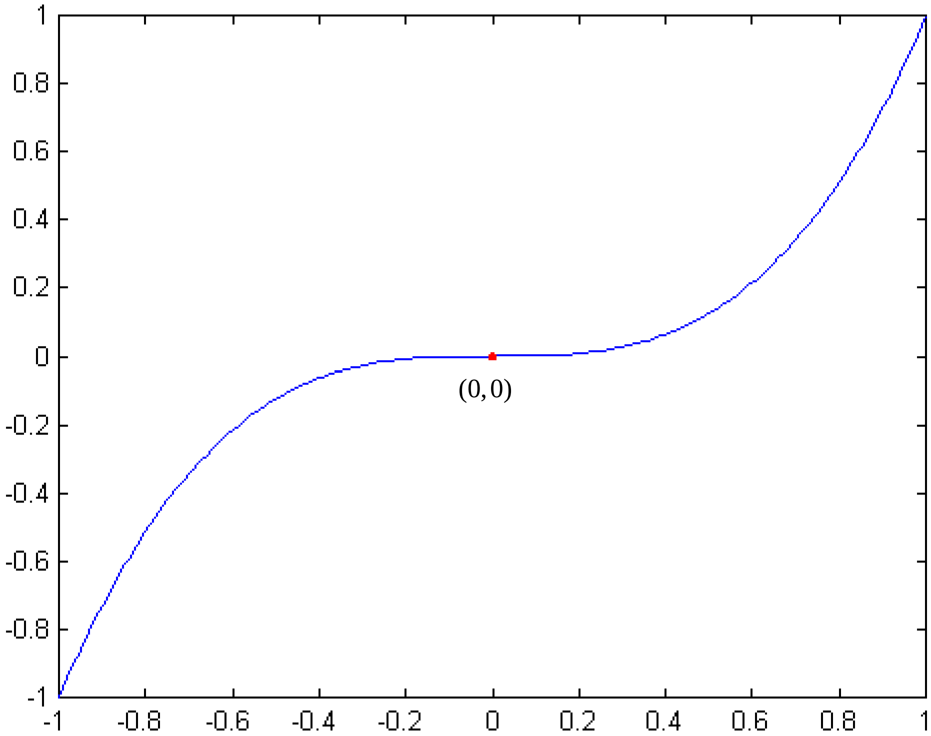

Figure 6: A graph of $f(x) = x^3$

Anyway, all these should give you a feeling that differentiation, which is closely related to slopes of graphs and rates of change, should also be very helpful in finding the local maximum and minimum of differentiable functions.

Exercise

You are highly advised to go through the following exercises to check that you have firmly grasp the right concepts. You should always state clearly where you have used the propositions and theorem mentioned above.

4.1 For each of the following functions, determine at what values of x is the function differentiable and find the derivatives of the functions at such x. Interpret the derivatives you obtain as slopes and rates of change.

(a) $f(x) = x^7 - 3x + 2$

(b) $g(x) = (x+1)(2x^2 - x^3)$

(c) $h(x) = \dfrac{x+2}{x^2+3}$

(d) $u(x) = \begin{cases} \dfrac{x^2-1}{x-1} + 1 & \text{if } x \neq 1 \\ 1 & \text{if } x = 1 \end{cases}$

(e) $v(x) = \begin{cases} x^2 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$

(f) $w(x) = |x|$

(g) $p(x) = |x^3|$

4.2 In Proposition 4.1.2 we proved that for all positive integers n, the function f defined by $f(x) = x^n$ is differentiable and $f'(x) = nx^{n-1}$ for all x where f is defined. Prove this result actually holds for all non-positive integers n as well. (Note that we say a function is differentiable if it is differentiable at all x where it is defined.) (Hint: use Proposition 4.2)

4.3 Find the local maximum(s) and minimum(s) of the following functions:

(a) $f(x) = x^3 - 5x^2 + x + 1$

(b) $g(x) = x^4 - x^2 + 7x$