Differentiation from First Principles

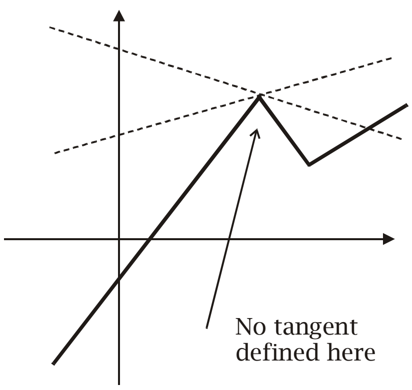
Prerequisites

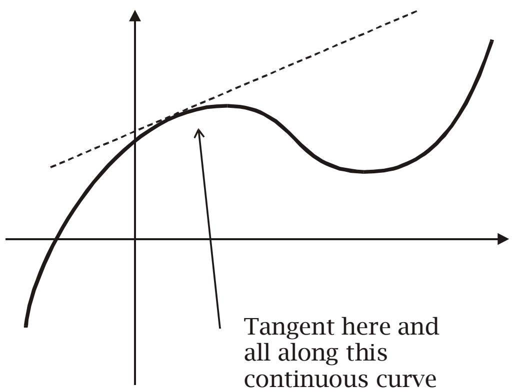
In this chapter it is assumed that you are familiar with all the standard derivatives of polynomial, trigonometric, exponential and logarithmic functions; likewise, it is assumed that you are familiar with all the basic rules of differentiation – for example, the Leibniz and chain rules for the differentiation of products and composites of function. You should be familiar with stationary points and how to test for them, both by means of the second derivative and by looking at how the sign of the derivative changes around the stationary point.

When you were first introduced to the differential calculus you may have learnt that it arises from the idea of trying to find gradients of functions, or tangents of their graphs. If that ever happened, it may have been now in the misty past, and the purpose of this chapter is to once again make you aware of the logical foundation of the calculus. We are going to call this *differentiation from first principles*, though as we shall remark at the end, the theory that is presented here itself only represents a stage on the way towards defining a real *foundation* for the calculus.

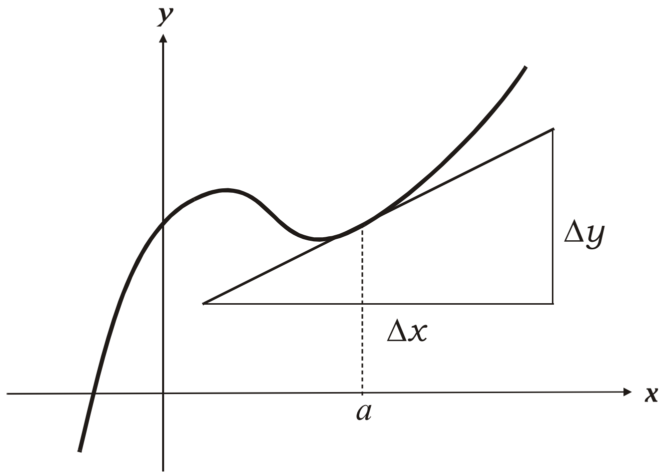
Tangent, Gradient and Rate of Change

A tangent is a line just touching a curve. Provided the curve has no sharp points there is just one tangent at every point.





The gradient of the tangent at *a* of a function *f(x)* represents the instantaneous rate of change of that function at *a*.

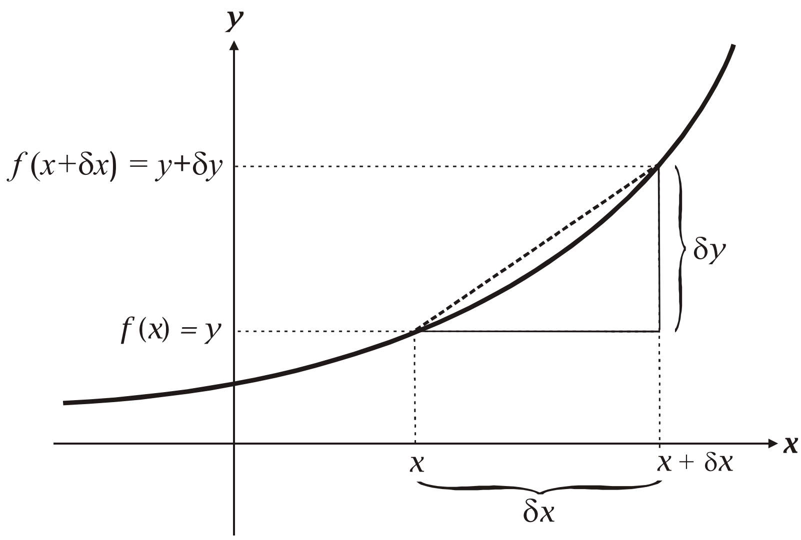


As the diagram indicates, the gradient can be found by graphical means.

$$\text{gradient} = m = \frac{\Delta y}{\Delta x} = \frac{\text{Change in } y}{\text{Change in } x}$$

Differential Calculus

The differential calculus provides a mathematical means of arriving at the gradient of a tangent to a function. Firstly, we calculate the gradient of a cord joining two points on the graph of the function:



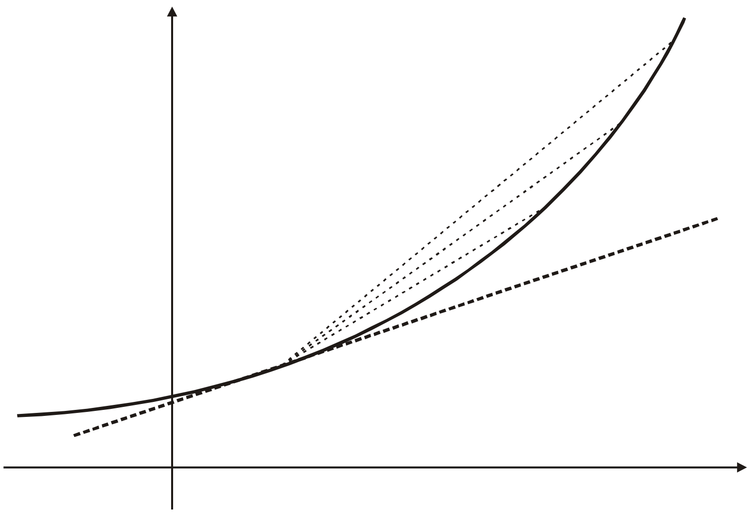
The gradient of the cord joining (x,y) to $(x+\delta x, y+\delta y)$ where δx and δy are small increments in *x* and *y* respectively, is

gradient = $\dfrac{\delta y}{\delta x} = \dfrac{f(x+\delta x)-f(x)}{\delta x}$

Sometimes the symbol *h* (or some other letter) is used for the small increment; in this case the gradient is given by

gradient = $\dfrac{\delta y}{\delta x} = \dfrac{f(x+h)-f(x)}{h}$

It is the same idea, just with a different letter. The differential calculus arises from the following idea. As δx gets smaller and smaller, the cord joining (x,y) to $(x+\delta x, y+\delta y)$ gets closer and closer to the tangent at *x*. The following diagram should help you to understand this.

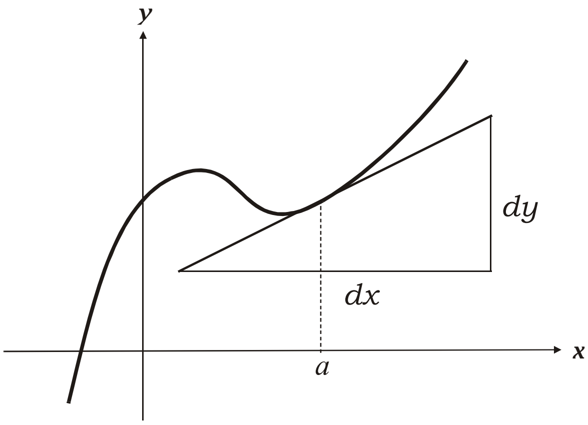


If δx becomes "infinitesimally small", then the cord becomes the tangent. Of course, there is a problem with the expression "infinitesimally small" – does it really mean anything. Actually, what we really mean is that as $\delta x \to 0$ (as δx gets closer and closer to 0) then the gradient of the cord gets closer and closer to the gradient of the tangent. This is really a statement about *limits* so eventually we have to recast the whole theory of the differential calculus in terms of limits. However, for now we will pretend that the phrase "infinitesimally small" does mean something, and also that in practice is amounts to putting $\delta x = 0$ at some point in the argument. We denote the gradient of the tangent by

$\dfrac{dy}{dx}$ or $f'(x)$ or y'

(There are many different symbols used for the derivative.)

Here dy and dx in $\frac{dy}{dx}$ mean the increments in y and x respectively along the tangent at a given point a to the graph of $y = f(x)$



Hence

$$\frac{dy}{dx} = \lim_{\delta x \to 0}\left\{\frac{\delta y}{\delta x}\right\} = \lim_{\delta x \to 0}\left\{\frac{f(x+\delta x) - f(x)}{\delta x}\right\} \qquad \text{or} \qquad \frac{dy}{dx} = \lim_{h \to 0}\left\{\frac{\delta y}{h}\right\} = \lim_{h \to 0}\left\{\frac{f(x+h) - f(x)}{h}\right\}$$

It is this definition that we call *the definition of the derivative from first principles.* Note here that δx and h are totally interchangeable. They mean the same thing, but we give both forms of the same equation because you see both in the literature. When you are asked (in this context) to find a derivative from first principles, it is with this equation (or equivalent) that you start.

Example (1)

The derivative of $y = f(x) = x^2$ from first principles is

$$\begin{aligned}
\frac{dy}{dx} &= \frac{d(x^2)}{dx} \\
&= \lim_{\delta x \to 0}\left\{\frac{f(x+\delta x) - f(x)}{\delta x}\right\} \\
&= \lim_{\delta x \to 0}\left\{\frac{(x+\delta x)^2 - x^2}{\delta x}\right\} \\
&= \lim_{\delta x \to 0}\left\{\frac{x^2 + 2x(\delta x) + (\delta x)^2 - x^2}{\delta x}\right\} \\
&= \lim_{\delta x \to 0}\left\{\frac{2x(\delta x) + (\delta x)^2}{\delta x}\right\} \\
&= \lim_{\delta x \to 0}\left\{2x + \delta x\right\} \\
&= 2x
\end{aligned}$$

$$\begin{aligned}
\frac{dy}{dx} &= \frac{d(x^2)}{dx} \\
&= \lim_{h \to 0}\left\{\frac{f(x+h) - f(x)}{h}\right\} \\
&= \lim_{h \to 0}\left\{\frac{(x+h)^2 - x^2}{h}\right\} \\
&= \lim_{h \to 0}\left\{\frac{x^2 + 2xh + h^2 - x^2}{h}\right\} \\
&= \lim_{h \to 0}\left\{\frac{2xh + h^2}{h}\right\} \\
&= \lim_{h \to 0}\left\{2x + h\right\} \\
&= 2x
\end{aligned}$$

Let us annotate this solution. At the line

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

the function $y = f(x) = x^2$ is substitute for *y*. This makes the next line strictly redundant. The next line

$$\frac{dy}{dx} = \lim_{\delta x \to 0}\left\{\frac{\delta y}{\delta x}\right\} = \lim_{\delta x \to 0}\left\{\frac{f(x+\delta x) - f(x)}{\delta x}\right\}$$

is the definition of the derivative from first principles. This is where we start, and actually the first line could be omitted. The function is $y = f(x) = x^2$ so we square $(x + \delta x)$ and x to obtain the next line

$$\frac{dy}{dx} = \lim_{\delta x \to 0}\left\{\frac{(x+\delta x)^2 - x^2}{\delta x}\right\}$$

This has removed the general function *f* and replaced it by the specific function $f(x) = x^2$. Now follows some algebra in which the term $(x + \delta x)^2$ is unsquared. This makes it possible to cancel out some terms. Then δx is a factor of both the numerator (top) and denominator (bottom) of the fraction, so it is cancelled through. At this point we have

$$\frac{dy}{dx} = \lim_{\delta x \to 0}\left\{2x + \delta x\right\}$$

Here we are allowed by our assumption that taking the limit is equivalent to putting $\delta x = 0$ so when we "*take the limit*" we merely substitute $\delta x = 0$ and scratch out the $\lim_{\delta x \to 0}$ part. Note, this step is only allowed if there is no zero in the denominator (the bottom) of a fraction, for otherwise it is equivalent to dividing by zero, and that is not allowed – it leads to contradictions.

Example (2)

Find the derivative of $\frac{1}{2x^2 - 3}$ from first principles.

Solution

This time we will give the solution just using the *h* notation.

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

$$= \lim_{h \to 0}\left\{\frac{\left\{\frac{1}{2(x+h)^2 - 3}\right\} - \left\{\frac{1}{2x^2 - 3}\right\}}{h}\right\} \qquad \left[\text{Using } f(x) = \frac{1}{2x^2 - 3},\ f(x+h) = \frac{1}{2(x+h)^2 - 3}\right]$$

$$
\begin{aligned}
\frac{dy}{dx} &= \lim_{h\to 0}\left\{\frac{\left\{\dfrac{1}{2x^2+4xh+h^2-3}\right\}-\left\{\dfrac{1}{2x^2-3}\right\}}{h}\right\} && [\text{Expanding the bracket}] \\
&= \lim_{h\to 0}\left\{\frac{\dfrac{2x^2-3-\left(2x^2+4xh+h^2-3\right)}{\left(2x^2+4xh+h^2-3\right)\left(2x^2-3\right)}}{h}\right\} && [\text{Putting the top over its common denominator}] \\
&= \lim_{h\to 0}\left\{\frac{-4xh-h^2}{h\left(2x^2+4xh+h^2-3\right)\left(2x^2-3\right)}\right\} && [\text{Cancelling out some terms, and tidying up}] \\
&= \lim_{h\to 0}\left\{\frac{-4x-h}{\left(2x^2+4xh+h^2-3\right)\left(2x^2-3\right)}\right\} && [\text{Dividing through top and bottom by } h] \\
&= -\frac{4x}{\left(2x^2-3\right)\left(2x^2-3\right)} && [\text{Taking the limit; i.e. substituting } h=0] \\
&= -\frac{4x}{\left(2x^2-3\right)^2} && [\text{Tidying up}]
\end{aligned}
$$

Note that in the expression $\dfrac{4x+h}{\left(2x^2+4xh+h^2-3\right)\left(2x^2-3\right)}$ we are allowed to substitute $h=0$ because the result does not lead to a zero on the bottom of the fraction. However, it is because of the potential threat of dividing by zeros that strictly speaking the whole calculus should be recast in terms of a theory of limits. But this is left for a later chapter.