**Parabola Volume 46, Issue 1 (2010)**

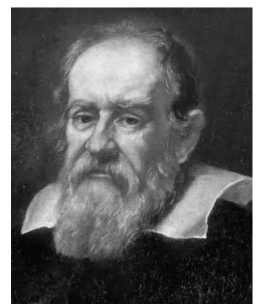
# History of the Derivative

**Milan Pahor**[1]

*Measure what is measurable, and make measurable what is not so.*
*Galileo Galilei*

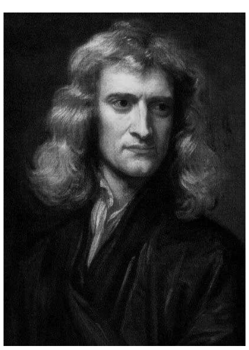
The 17th century was a revolutionary period in the development of modern science. In 1610 Galileo Galilei began a detailed study of the moons of Jupiter using the cutting edge technology of the telescope. By 1619 Johannes Kepler (implementing the detailed data he inherited after Tycho Brahe's death) had published his three laws of planetary motion, the first of which made the inflammatory assertion that the orbit of each planet in the solar system is an ellipse with the sun at one focus.

Scientists (as natural philosophers became known) began to model the physical world through the emerging techniques of experimentation and mathematical analysis. Despite some spirited opposition (in 1633 Galileo was brought before the Inquisition, shown the various instruments of torture and invited to recant, which he did) the end of the century saw Isaac Newton laying the foundations of modern science with his laws of motion and gravitation.



Galileo
(1564-1642)

Mathematics as it stood was effective in dealing with fixed well defined static objects. However it soon became clear that this developing scientific analysis of the physical world demanded an entirely new theory of mathematics capable of dealing with evolving systems in a state of flux. Against this backdrop, the fundamental concepts of differential calculus began to surface across Europe. Ultimately the threads of the theory were brought together by Sir Isaac Newton (1642-1727) in England and independently by Gottfried Wilhelm von Leibniz (1646-1716) in Germany.



Newton
(1642-1727)

[1] Milan Pahor is a Lecturer in Mathematics and Statistics at the University of New South Wales.

To Newton a curve was simply the path of a moving particle. From his point of view the essential problem of calculus was the subsequent calculation of its velocity. In other words he needed to be able to jump from knowledge of where a particle was, to knowledge of where it was heading. Newton viewed calculus as an essential tool in his analysis of mechanics and gravitation.

Leibniz
(1646-1716)

The discovery of calculus is unfortunately mired in controversy (on many different levels) with Leibniz also staking a claim to its discovery. His approach was more abstract in that he simply sought to establish the gradient of the tangent at any point on a curve. Leibniz was also one of the first to view a curve as the graph of an algebraic relation and to couch the derivative in terms of the theory of functions.

From a modern perspective both of these problems are equivalent to the calculation of the derivative $\dfrac{dy}{dx}$, the instantaneous rate of change of a function $y = f(x)$.

Both men spent their later years locked in a bitter dispute over ownership. Due in some part to the enormous power he wielded as the President of the Royal Society, Newton seems to have secured the historical credit as the father of calculus. However there is little doubt that Leibniz's approach and notation, which has survived to the present day, offers a deeper insight into the fundamental workings of the derivative.

Although these terms were not in use until much later, Leibniz's abstract, almost philosophical, approach to the derivative was that of the Pure mathematician whereas the pragmatic Newton could be described as an Applied mathematician or a Physicist. It is fair to say however that neither man fully resolved the technicalities surrounding the formal definition of the derivative. That would take another 150 years of intense analysis. We will examine the two different techniques of Newton and Leibniz in detail. But let us first take a look at a 2009 2 unit HSC question as a foil.

**2009 Higher School Certificate**
**(2 unit Mathematics paper Question 1d)**

**Question:** Find the gradient of the tangent to the curve $y = x^4 - 3x$ at the point $(1, -2)$.

The concepts which surround this problem are extremely subtle. Note firstly that this question could be asked in a number of different ways.

**Question 2:** Find the gradient of $y = x^4 - 3x$ at the point $(1, -2)$.

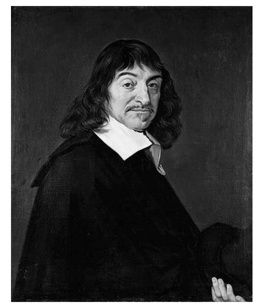
**Question 3:** Find the instantaneous rate of change of $y = x^4 - 3x$ at the point $(1, -2)$.

**Question 4:** For $y = x^4 - 3x$ find $\dfrac{dy}{dx}$ at the point $(1, -2)$.

The following thoughts will be swirling through a student's mind. Given a function $y = f(x)$ there exists another function called the derivative or the gradient function and denoted by $\dfrac{dy}{dx}$ or $f'(x)$. This gradient function $\dfrac{dy}{dx}$ has the property that its value $f'(1)$ at $x = 1$ is the gradient of the original function at $x = 1$. I have a formal definition for $\dfrac{dy}{dx}$ but don't need to use it because I also have a bag of algorithms, facts and tricks that I can use to find the derivative. The process of differentiation is linear so I can bust the problem up into smaller bits. I also have the simple formula $\dfrac{d}{dx}x^n = nx^{n-1}$. So...

$$f'(x) = 4x^3 - 3 \rightarrow f'(1) = 4 \times 1 - 3 = 1.$$

What is especially revealing is the position of this question on the paper, Question 1(d). Thus the examiners consider this to be one of the simplest questions on the exam. But there is 300 years of blood, sweat, tears and heartache in the above analysis! Let's step back 350 years and take a look at how Rene Descartes (1596-1650) attacked such a problem in the period just before the development of the derivative.

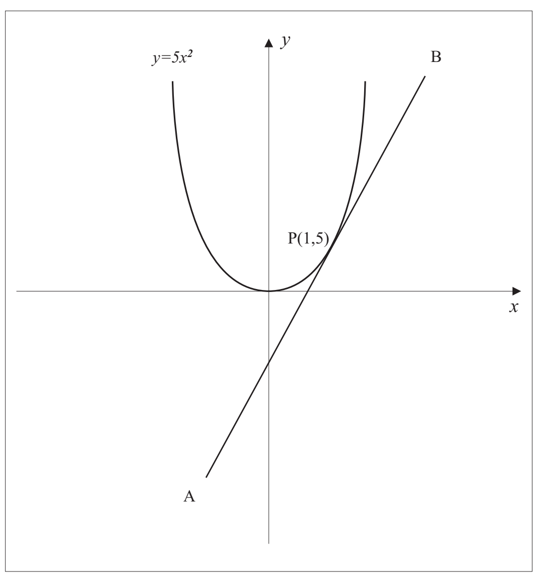


Descartes (1596-1650)

We will be testing out all of the different approaches to differentiation on the function $y = 5x^2$.

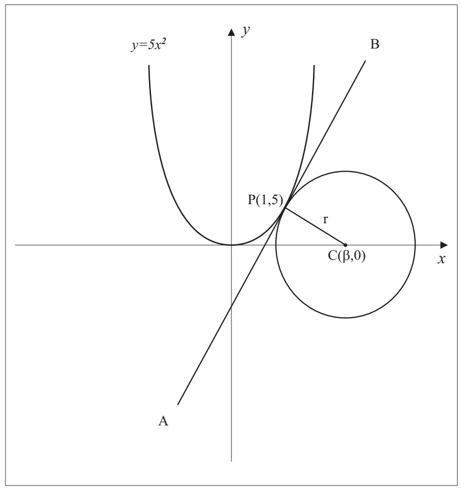
**Descartes' Approach**

**Example 1:** Find the gradient of the tangent to $y = 5x^2$ at $x = 1$.




The first observation that Descartes would have made is that this is a tough problem. Everything moves! As the point $P$ slides around on the curve the tangent $AB$ is constantly changing, as is its gradient. Secondly the calculation of the gradient of a curve is a crucial technique to have at ones' disposal. The function itself controls the fixed value of $y$, but the gradient is a measure of its evolution!

Descartes wrote that the problem of constructing tangent lines is 'the most useful and general problem that I know or even have ever desired to know in geometry'. This is how he did it:




We construct a circle with centre $(\beta, 0)$ on the $x$-axis and radius $r$ chosen so that the circle and the parabola have a common tangent $AB$ at $P$. The equation of the circle is then $(x-\beta)^2 + y^2 = r^2$ and intersecting the circle with the parabola yields $(x-\beta)^2 + (5x^2)^2 = r^2$ and hence $(x-\beta)^2 + 25x^4 = r^2$. Upon expanding we have $x^2 - 2\beta x + \beta^2 + 25x^4 = r^2$ and hence

$$25x^4 + x^2 - 2\beta x + (\beta^2 - r^2) = 0.$$

Now this equation has a double root at $x = 1$ so for some $b$ and $c$

$$\begin{aligned} 25x^4 + x^2 - 2\beta x + (\beta^2 - r^2) &= (x-1)^2(25x^2 + bx + c) \\ &= (x^2 - 2x + 1)(25x^2 + bx + c) \\ &= 25x^4 + (b-50)x^3 + (c - 2b + 25)x^2 \\ &\quad +(b-2c)x + c \end{aligned}$$

Equating powers of $x$ we have

$b - 50 = 0 \to b = 50$

$c - 2b + 25 = 1 \to c = 76$

$-2\beta = b - 2c = -102 \to \beta = 51.$

Now $m_{PC} = \dfrac{5-0}{1-\beta} = \dfrac{5-0}{1-51} = -\dfrac{1}{10}$. Since $PC \perp AB$ we can use $m_1 m_2 = -1$ to obtain

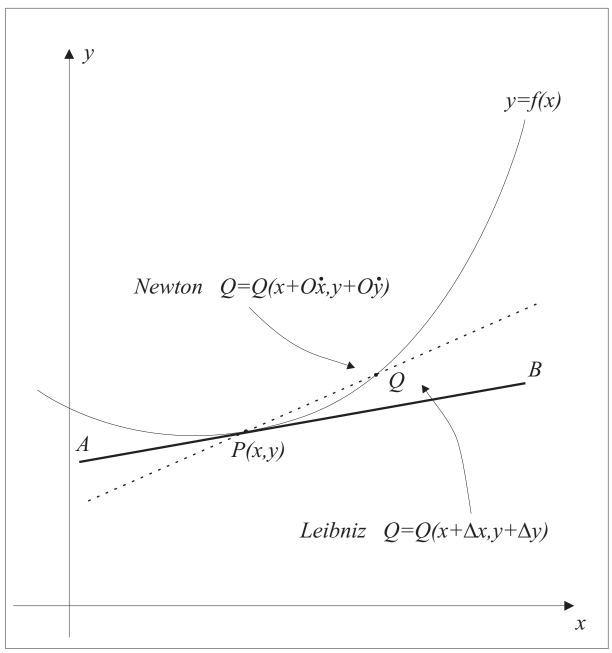
$$m_{AB} = 10.$$

The contrast with a modern approach is alarming:

$$f(x) = 5x^2 \to f'(x) = 10x \to f'(1) = 10.$$

Clearly Descartes' method has major problems. Even in the quadratic case the algebra is ponderous and the technique simply did not have the legs to be of any genuine use to Newton. Nevertheless it does sort of work. It is easy to sit back and have a bit of a chuckle when comparing the two methods above, however keep in mind that Rene Descartes was one of the greatest thinkers of his age.

We now contrast the techniques of Newton and Leibniz on the simple question of finding the gradient of a tangent to a curve $y = f(x)$. Both approaches begin by choosing a point $P$ on the curve where the tangent is to be placed. A second point $Q$ is then positioned on the curve 'near $P$'. The heart of the difference between the two approaches is the definition and perception of $Q$. The tangent $AB$ is then analysed by instead considering the gradient of the secant $PQ$ as $Q$ approaches $P$ in a vague undefined manner.




**Newton's Approach**

Newton's approach is all about time and motion. He considered the curve to be the path of a moving particle and referred to the variables $x$ and $y$ as fluents (they flowed). The time rates of change $\dot{x}$ and $\dot{y}$ were then named fluxions. Suppose that the particle was initially at a point $P(x, y)$ and consider its position on the curve after a small interval of time $O$ had elapsed. Since $distance = time \times velocity$ the particle will have moved to $Q(x + O\dot{x}, y + O\dot{y})$. Clearly $Q$ is still on the curve and hence must satisfy its equation $y = 5x^2$. Thus

$$y + O\dot{y} = 5(x + O\dot{x})^2$$

$$\begin{aligned} &= 5(x^2 + 2xO\dot{x} + O^2\dot{x}^2) \\ &= 5x^2 + 10xO\dot{x} + 5O^2\dot{x}^2 \\ c &= y + 10xO\dot{x} + 5O^2\dot{x}^2 \end{aligned}$$

So $O\dot{y} = 10xO\dot{x} + 5O^2\dot{x}^2$ and division by the small time interval $O$ yields $\dot{y} = 10x\dot{x} + 5O\dot{x}^2$. Letting $O$ disappear ($Q \to P$) we have $\dot{y} = 10x\dot{x}$ which implies that

$$\frac{\dot{y}}{\dot{x}} = 10x.$$

A modern closure would be that $\frac{\dot{y}}{\dot{x}} = \frac{\frac{dy}{dt}}{\frac{dx}{dt}} = \frac{dy}{dt}\frac{dt}{dx} = \frac{dy}{dx}$.

However Newton was not interested in $\frac{dy}{dx}$ ! From his point of view the quantity $\frac{\dot{y}}{\dot{x}}$ was *by definition* the gradient of the tangent. In the language of the day this was called a solution to the tangent problem.

**Leibniz's Approach**

Leibniz made no use of time at all and instead dealt directly with the function $y = 5x^2$. He simply moved $P(x, y)$ through small increments $\Delta x$ and $\Delta y$ to $Q(x + \Delta x, y + \Delta y)$. Then

$$y + \Delta y = 5(x + \Delta x)^2 = 5(x^2 + 2x\Delta x + (\Delta x)^2) = 5x^2 + 10x\Delta x + 5(\Delta x)^2$$

Since $y = 5x^2$ we have $\Delta y = 10x\Delta x + 5(\Delta x)^2 \to \frac{\Delta y}{\Delta x} = 10x + 5\Delta x$. Letting the infinitesimal $\Delta x$ vanish ($Q \to P$) we have the gradient of the tangent being $10x$.

There is very little difference between the two approaches. Unfortunately Newton was an extremely secretive character. He was reluctant to publish and even when he did make his results known, he would disguise the material through various codes. Leibniz on the other hand wanted to tell the world. It is clear from various letters that Newton had beaten Leibniz to the discovery of calculus by some 10 years but it was certainly Leibniz who published first.

**The Problem of Infinitesimals**

Both arguments suffer from the same fatal flaw. Neither man had the slightest notion as to a formal definition of a limit. Newton had his infinitesimal time interval $O$ popping in and out of existence. Likewise Leibniz had the infinitesimal $\Delta x$ mysteriously disappearing and reappearing. These infinitesimals were non-zero when they needed to be non-zero (for division) but suddenly became zero to close the argument.

The reality of infinitesimals became a significant philosophical problem. It seems the case that both Newton and Leibniz viewed infinitesimals simply as a convenient notation. However many of their colleagues, the Bernoulli brothers in particular, were adamant that infinitesimals actually existed lurking somewhere between zero and the start of the positive real numbers. This explained their capacity to exist as both zero and non-zero quantities.

Not surprisingly, the possibility of objects springing in and out of existence aroused the interest of the church. Bishop Berkeley (quite rightly) launched a stinging attack on the theory:

'and what are these fluxions? The velocities of evanescent increments. And what are these same evanescent increments? They are neither finite quantities, nor quantities infinitely small, nor yet nothing. May we not call them ghosts of departed quantities?'

It would be a century and a half before the derivative had a formal limit definition.

Interestingly some recent effort has gone into what is called the hyperreal number system. These include the usual real numbers together with actual living infinitesimals, that is non zero 'numbers' whose absolute value is smaller than that of any real number. Remarkably it is possible to logically reformulate the derivative in terms of these infinitesimals rather than through limits.....but it's pretty spooky.

**Notation**

One issue that does seem to be agreed upon is that Leibniz's notation is significantly superior to that of Newton. For example:

(i) Suppose that $y = f(u)$ where $u = g(x)$.

The chain rule in Newton's terminology is $y' = f'(u)u'$ while in Leibniz's terms it is

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}.$$

Leibniz's formula gives the reader a clear feeling as to why it all works! The applications of Leibniz's notation are sometimes almost miraculous:

(ii) Suppose that $y$ is a product of two quantities $y = uv$.

We make $y$ a little larger

$y^+ = (u + \frac{1}{2}\Delta u)(v + \frac{1}{2}\Delta v) = uv + \frac{1}{2}u\Delta v + \frac{1}{2}v\Delta u + \frac{1}{4}\Delta u\Delta v$

and then a little smaller

$y^- = (u - \frac{1}{2}\Delta u)(v - \frac{1}{2}\Delta v) = uv - \frac{1}{2}u\Delta v - \frac{1}{2}v\Delta u + \frac{1}{4}\Delta u\Delta v$

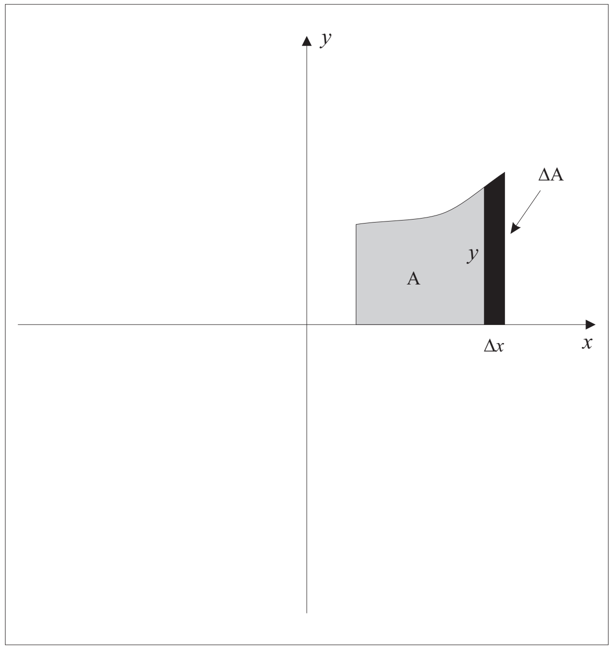
Subtracting we have

$y^+ - y^- = \Delta y = u\Delta v + v\Delta u$ and dividing by the infinitesimal $\Delta x$ yields

$$\frac{\Delta y}{\Delta x} = u\frac{\Delta v}{\Delta x} + v\frac{\Delta u}{\Delta x}$$

The product rule!!

Both Newton and Leibniz were fully aware of the fundamental theorem of calculus and referred to it as the inverse tangent problem. The 'proof' using Leibniz's notation is delightfully simple.




Let $A$ be the area accumulated underneath the curve $y = f(x)$ and consider the extra increment of area $\Delta A$. This is approximately a rectangle so $\Delta A = y\Delta x$. Hence $\dfrac{\Delta A}{\Delta x} = y$. It follows that we need to 'antidifferentiate' $y$ to find $A$.

**Limits**

The controversy and struggle over infinitesimals took almost 150 years to resolve. In the 1800s Augustin-Louis Cauchy (1789-1857) and Karl Weierstrass (1815-1897) established the formal definition of the limit of a function $\lim_{x\to a} f(x)$.

The vaguely mystical properties of infinitesimals were replaced with a brutally exact definition of the limit. We say that $\lim_{x\to a} f(x) = L$ if and only if:

For each $\epsilon > 0$ there exists $\delta > 0$ such that $|x - a| < \delta$ implies that $|f(x) - L| < \epsilon$.

Intuitively this means that $f(x)$ can be made as close to $L$ as you wish by simply choosing $x$ sufficiently close to $a$.

Weierstrass
(1815-1897)

It is crucial when discussing limits with students not to fall into the same traps as Newton and Leibniz. If $\lim_{x \to a} f(x) = L$ do not say that the limit approaches $L$, or that the limit is approximately equal to $L$.........the limit *is* $L$!

Cauchy
(1789-1857)

It is fair to say that most undergraduate students struggle to come to grips with the $\epsilon - \delta$ definition of a limit. But the definition is completely bullet-proof.

The door was now open for our modern precise formal definition of the derivative.

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

Returning to our motivating example of $y = f(x) = 5x^2$ the story comes to a close:

$$\begin{aligned} \frac{dy}{dx} &= \lim_{h \to 0} \frac{5(x+h)^2 - 5x^2}{h} \\ &= \lim_{h \to 0} \frac{5(x^2 + 2xh + h^2) - 5x^2}{h} \\ &= \lim_{h \to 0} \frac{5x^2 + 10xh + 5h^2 - 5x^2}{h} \\ &= \lim_{h \to 0} \frac{10xh + 5h^2}{h} \\ &= \lim_{h \to 0} 10x + 5h \\ &= 10x. \end{aligned}$$

The formal definition of the derivative is the culmination of 200 years of intense mathematical development. Just as an advanced English student would never be allowed to leave school without exposure to Shakespeare, a student of calculus must have mastered differentiation using limits.

What became of Gottfried Wilhelm von Leibniz? He was completely crushed by the immensely powerful Newton. Leibniz was happy to share the glory but Newton was adamant that his work had been stolen. Leibniz became increasing marginalised and found it difficult to secure academic positions. When he died in 1716 his funeral was attended only by his personal secretary and his grave remained unmarked for almost 50 years.

Newton wrote 'If I have seen further it is by standing on the shoulders of Giants'.

He also trod on a few toes.

**Acknowledgement**

The figures for this article were sourced from Wikimedia Commons:
http://www.wikipedia.org/