THE DERIVATIVE

Esperanza Blancaflor Arugay

In the seventeenth century, an Englishman, Sir Isaac Newton (1642-1727), and a German, Gotfried Wilhelm Leibnitz (1646-1716), independently introduced the calculus. The calculus is considered the motion-picture machine of mathematics. It catches natural phenomena in the act of changing, or as Newton called it, in a state of flux. Other fields of mathematics are to be likened to the camera which shows a still picture of nature as it appears at a given instant without regard to the possible appearance the following instant.

Nature is never static. Everything around us is in a state of motion. Hence, the processes of nature is not possible without the notion of rate of change or derivative.

Rate of change: the changing time of sunset. The sun sets at different times of the year, depending on date and location. Here in the Philippines, in the year 1994, for instance, the sun sets at

6:14 on May 5,

5:43 on October 5, and

5:25 on December 5.

The time of sunset is a *function* of the date. If we let T be the time of sunset in hours and minutes and d be the date of the year, we can express our functional relation in the form $T = T(d)$. For example, from the data above, we have 6:14 = T(May 5), or we simply write 6:14 = $T(125)$, since May 5 is the 125th day of the year. Similarly, we can write 5:43 = $T(278)$ and 5:25 = $T(339)$. We observe that the rate at which the time of sunset is changing varies at different times of the year. We show how the rate varies by looking at some further data for sunset from a weather news for the year 1994.

Using the data below, we can estimate the rate at which the time of sunset is changing on October 5. On October 8, the sun sets 5 minutes earlier than on October 2. This is a change of –5 minutes in six days, so the rate of change is

$$(-5 \text{ minutes }) / (6 \text{ days}) \approx -.83 \text{ minutes per day.}$$

date	time	date	time	date	time
May 2	6:13	October 2	5:45	Dec. 2	5:25
May 5	6:14	October 5	5:43	Dec. 5	5:25
May 8	6:15	October 8	5:40	Dec. 6	5:26

We say this is the rate at which sunset is changing on October 5 and we write,

$$T'(278) \approx -.83 \text{ minutes per day.}$$

The negative sign indicates that the time of sunset is decreasing, i.e., the sun is setting earlier each day. Similarly, we find that around December 5, $T'(339) \approx .17$ minutes per day and around May 5, $T'(125) \approx .33$ minutes per day. The last two values are positive since the time of sunset is increasing, i.e., the sun is setting later each day in May and December.

With these rates, we can estimate the time of sunset for dates not given in the table. For instance, May 10 is five days after May 5, so the total change in the time of sunset from May 5 to May 10 should be approximately

$$\Delta T \approx (.33 \text{ min. per day}) \times -5 \text{ days} = 1.65 \text{ minutes.}$$

In whole numbers, then, the sun sets 2 minutes later on May 10 than on May 5. Since sunset on May 5 is 6:14, sunset on May 10 is 6:16.

By letting the change in number of days be negative, we can use the same reasoning to tell us the time of sunset on days shortly before the given dates. For example, November 29 is –6 days away from December 5, so the change in the time of sunset should be

$$\Delta T \approx (.17 \text{ min. per day}) \times -6 \text{ days} = -1.02 \text{ minutes.}$$

Therefore, we can estimate that sunset occurred at 5:25 – 0:01 = 5:24 on November 29.

Changing rates. Suppose, instead of using the tabulated values for October 5, we try to use our May data to predict the time of sunset in October 5. Now, October 5 is 153 days away after May 5, so the change in the time of sunset should be approximately

$$\Delta T \approx (.33 \text{ min. per day}) \times 153 \text{ days} = 50.49 \text{ minutes},$$

and we compute that sunset on October 5 should be 6:14 + 0:50 = 6:64 or 7:04, which is 1 hour and 21 minutes later than the actual time! When we use the formula above to estimate ΔT, we assume that the time of sunset changes at a *fixed rate* of .33 minutes per day for the entire 153-day time span. The rate actually varies, and the variation is too great for us to get a good estimate. Only with a much smaller time-span does the rate not vary too much. *Predictions over long time spans are less reliable.*

REMARK. While many rates do involve changes with respect to time, other rates do not. Examples are the dose rate for medicine (milligrams per pound of body weight), annual birth rate (live births per 1,000 population), death rates (deaths per 1,000 population). Any quantity expressed as a percentage, such as an interest or unemployment rate, is a rate of similar sorts. An unemployment rate of 4% , for instance, means 4 unemployed workers per 100 workers.

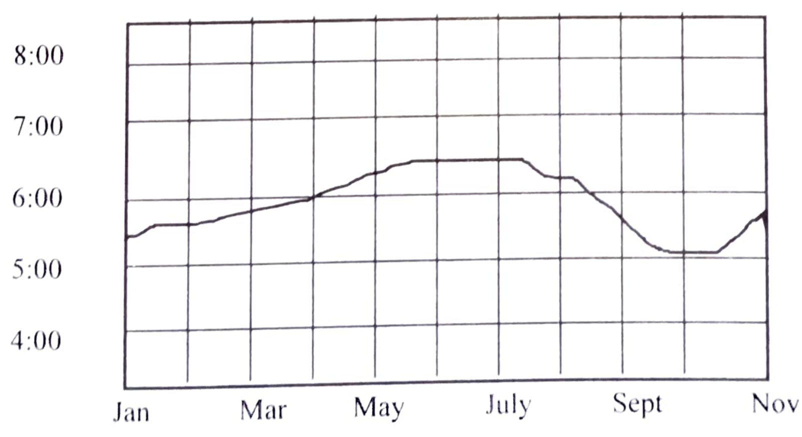

Figure 1

Graph of a function; graph of data. We now look at the graph of the sunset function. See Figure 1. The dates are represented in the horizontal axis, the times are on the vertical axis.

We know from geometry that the rate of change of a linear function can be visualized as the slope of its graph. Can we say the same thing about the sunset function?

Looking at the graph of the sunset function, we observe that its graph is not a line, so the sunset function is not linear. How do we make the connection then between rate and slope? What do we mean by the slope of this graph? These can be answered and made clear by enlarging the graph.

Imagine we have a "microscope" that allows us to "zoom in" on the graph near each of the three dates we considered in the given data. If we put each magnified image in a window, we get windows *a*, *b* and *c* below.

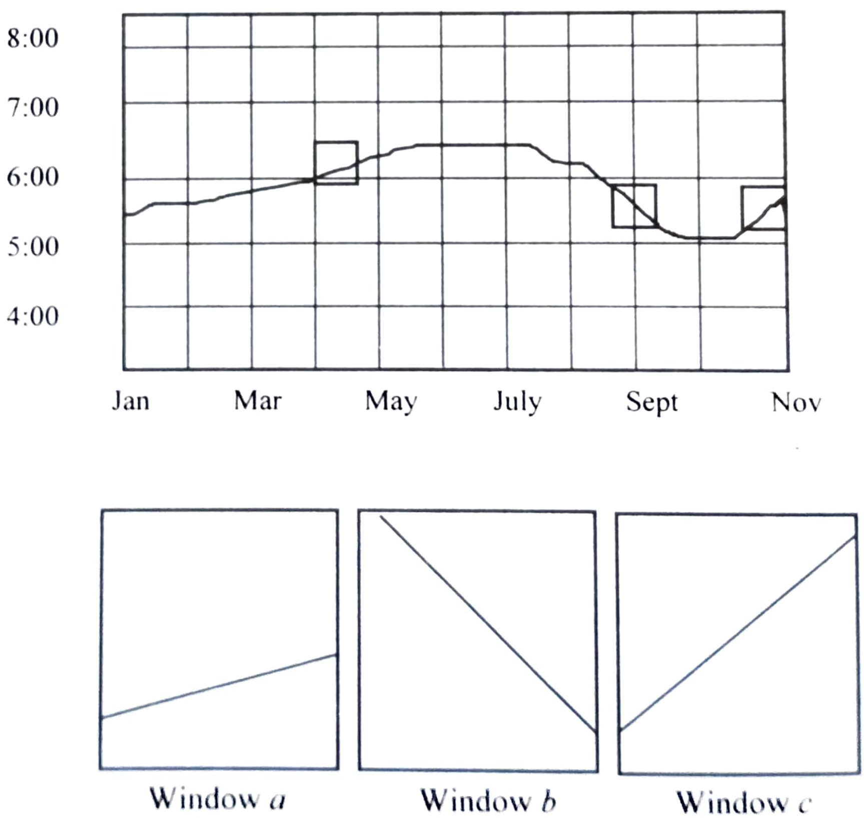

Figure 2

Notice how different the graph looks under the microscope. First of all, it now shows up clearly as a collection of separate points - one for each day of the year. Second, the points in a particular window lie on a line that is *essentially straight*. The straight lines in the three windows have very different slopes, but that is only to be expected.

Let us calculate the slope in each window. In window a, choose the points $(d_1, T_1) = (122, 6{:}13)$ and $(d_2, T_2) = (128, 6{:}15)$. The slope is

$$\frac{\Delta T}{\Delta d} = \frac{T_2 - T_1}{d_2 - d_1} = \frac{6{:}15 - 6{:}13}{128 - 122} = \frac{2 \text{ minutes}}{6 \text{ days}} \approx .33 \frac{\text{minutes}}{\text{day}}.$$

Using the same approach in the other two windows we find that the slope in window b is $-.83$ min/day and the slope of the line in window c is .17 min/day. These are exactly the same values as we have for rates of change of the time of sunset around May 5, October 5, and December 5.

REMARK. The rate of change of a function at a point is equal to the slope of its graph at that point, if the graph looks straight when we view it under a microscope.

The graph of a formula. Rates and slopes are really the same thing - that's what we learn by using a microscope to view the graph of the sunset function. But the graph of the sunset function consists of a finite number of disconnected points - this is a problem when we deal with data. In such cases, magnifying the graph too much becomes useless. For instance, we get no information from a window that was narrower than the space between the data points. On the other hand, if we consider the graph of a function given by a formula, there is no limitation of magnification. We can magnify as much as we wish and still obtain useful information.

While the microscope is used by biologists to study microorganisms, the "microscope" that we are talking about here is actually a computer in which we can enlarge or shrink any part of a graph by pressing a "zoom in" or "zoom out" command on the keyboard of the computer.

Consider the function $f(x) = x^4 - 8x$. Let's find $f'(1)$, the rate of change of f when $x = 1$. We need to zoom in on the graph of f at the point $(1, f(1)) = (1, -7)$. We do this in stages, producing a succession of windows (See Figure 3) that run clockwise from the upper left corner.

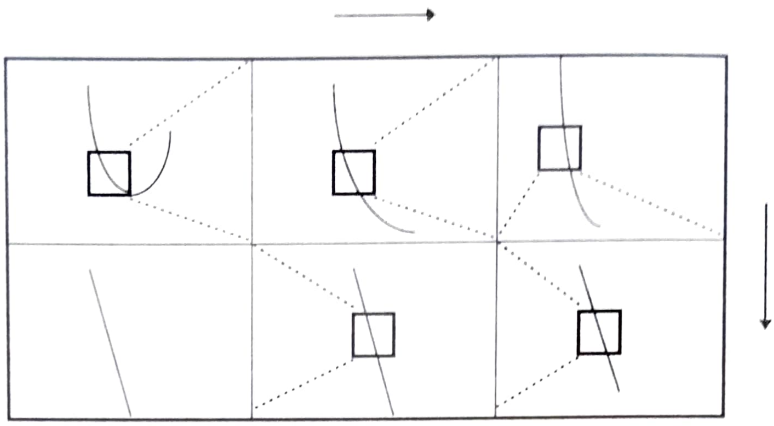

Figure 3

Notice how the graph gets straighter with each magnification. Let's call ***field of view*** that part of the graph that we see in a window. The field of view of the second window is only one-tenth as wide as the previous one, and the field of view of the last window is only one-thousandth of the first!

Intuitively, the rate of f is the slope of the graph of f when we magnify the graph enough to make it look straight. But how much is enough? Which window should we use? The following table gives the slope $\Delta y/\Delta x$ of the line that appears in each of the last four windows in the sequence. For Δx we take the difference between the x-coordinates of the points at the ends of the line, and for Δy we take the difference between the y-coordinates.

Δx	Δy	$\Delta y/\Delta x$
.1528	–.589638844	–3.858892958
.0147	–.058823280	–4.001583673
.0069	–.027603811	–4.000552319
.0037	–.014802170	–4.000586486

As you can see, under successive magnifications, the first five digits of $\Delta y/\Delta x$ have stabilized. The values of $\Delta y/\Delta x$ are successive approximations to the slope of the graph. The exact value of the slope is then the limit of these approximations as the width of the field of view shrinks to zero:

$$f'(1) = \text{the slope of the graph at } x = 1 = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}.$$

In the limit process, we take $\Delta x \to 0$ because Δx is the width of the field of view. Since five digits of $\Delta y/\Delta x$ have stabilized, we can write

$$f'(1) \approx 4.0005 .$$

To find $f'(x)$ at some other point, proceed the same way. Magnify the graph at that point repeatedly, until the value of the slope stabilizes. Our main observations are summarized below.

REMARKS. (1) The ***slope*** of a graph at a point is the limit of the slopes seen in a computer at that point, as the field of view shrinks to zero.

(2) The ***rate of change*** of a function at a point is the slope of its graph at the point. Thus the rate of change is also a limit.

Formal definition of the derivative. The word derivative is adopted for convenience to replace the more cumbersome phrase *rate of change*. In other words, if we can compute the derivative of a function $y = f(x)$ at $x = x_1$ and can find its value, the number thus found tells us how fast f is changing with respect to x at the point on the graph whose abscissa is x_1.

2.1 DEFINITION. The ***derivative of the function*** f at x is the rate of change of f at the point $(x, f(x))$, which is the same as the slope of its graph at $(x, f(x))$. This is denoted by $f'(x)$. Formally, it is defined by

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

The value $f'(x)$ is called the ***derivative of*** f ***at*** x, if the limit exists.

Let us examine this definition from the geometric point of view.

Consider the graph of a function $y = f(x)$. Let $P(x_1,y_1)$ be a point on the graph and $Q(x_2,y_2)$ be another point of the graph near P. The line through the points P and Q is called a secant line. (Figures 4 and 5).

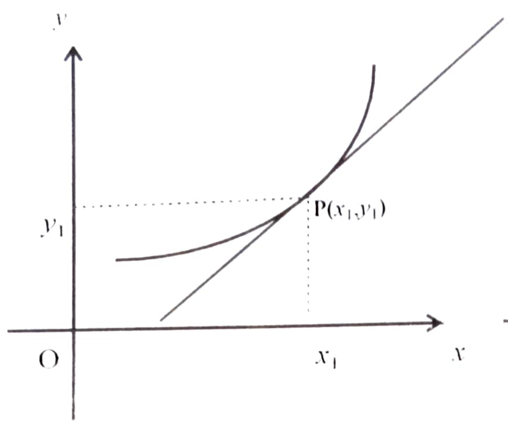

Figure 4

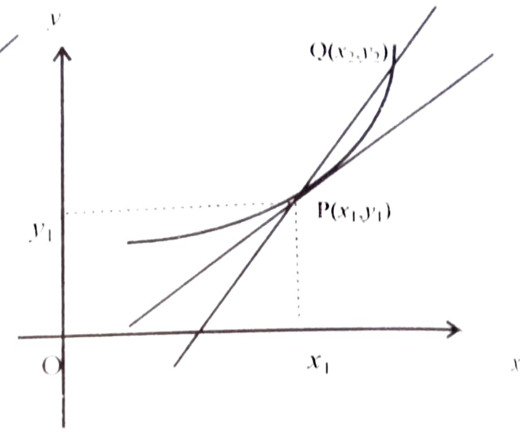

Figure 5

Note that $x_2 = x_1 + \Delta x$ and $y_2 = y_1 + \Delta y$. The slope of the secant line is,

$$\frac{y_2 - y_1}{x_2 - x_1} = \frac{\Delta y}{\Delta x} = \frac{f(x_1 + \Delta x) - f(x_1)}{\Delta x}$$

Now, if we let Δx approach 0, the point Q will move along the curve $y = f(x)$ and approach the point P. What happens to the secant line? As the point Q moves closer to the point P, the secant line pivots about the point P and gets closer to the tangent line (see Figure 6).

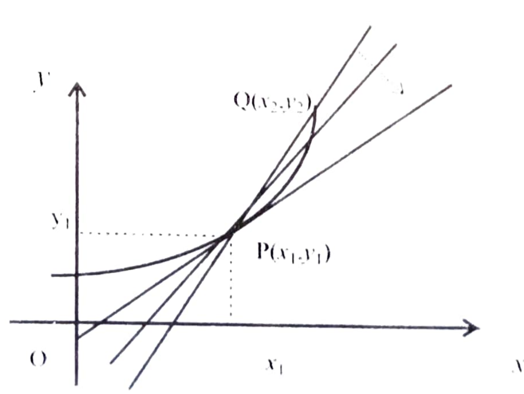

Figure 6

Thus as $\Delta x \to 0$, the slope $\Delta y/\Delta x$ of the secant line approaches the slope m of the tangent line, that is,

$$m = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x_1 + \Delta x) - f(x_1)}{\Delta x} = f'(x_1).$$

The geometric interpretation of the derivative above shows to us that the derivative of a function f at a point $x = x_1$ in the domain of f is the slope of the tangent line to the graph of f at the point $(x_1, y_1) = (x_1, f(x_1))$. The expression $\dfrac{f(x_1 + \Delta x) - f(x_1)}{\Delta x}$ is called a ***difference quotient.***

REMARK. The derivative is a limit. To find that limit, we follow a four step process:

(a) Replace y by $y + \Delta y$ and x by $x + \Delta x$ in the given equation.

(b) Solve for Δy.

(c) Divide both sides of the equation by Δx.

(d) Evaluate $\lim\limits_{\Delta x \to 0} \dfrac{\Delta y}{\Delta x}$.

2.1.1 EXAMPLE. Given $y = x^3$, find $f'(x)$.

Solution. Since $y + \Delta y = (x + \Delta x)^3$, we immediately have

$$\begin{aligned} \Delta y &= (x + \Delta x)^3 - y \\ &= (x^3 + 3x^2\Delta x + 3x(\Delta x)^2 + (\Delta x)^3 - x^3 \\ &= 3x^2\Delta x + 3x(\Delta x)^2 + (\Delta x)^3. \end{aligned}$$

$$\therefore \frac{\Delta y}{\Delta x} = \frac{3x^3\Delta x + 3x(\Delta x)^2 + (\Delta x)^3}{\Delta x}.$$

$$\therefore \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{3x^2\Delta x + 3x(\Delta x)^2 + (\Delta x)^3}{\Delta x} = 3x^2.$$

Hence, $f'(x) = 3x^2$. #

NOTE. In calculating this limit, you must be careful to treat x as a "constant" while letting Δx approach zero.

2.1.2. EXAMPLE. If $f(x) = \sqrt{x}$, find $f'(x)$.

Solution. From $f(x) = \sqrt{x}$, we have $f(x+\Delta x) = \sqrt{x+\Delta x}$. Hence,

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

$$= \lim_{\Delta x \to 0} \frac{\sqrt{x+\Delta x} - \sqrt{x}}{\Delta x} \frac{\sqrt{x+\Delta x} + \sqrt{x}}{\sqrt{x+\Delta x} + \sqrt{x}}$$

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

$$\therefore \; f'(x) = \frac{1}{2\sqrt{x}}. \quad \#$$

The preceding example makes use of the definition of the derivative in functional notation.

The derivative notations. Other notations used to denote the derivative of $y = f(x)$ are:

$$\frac{dy}{dx},\ y',\ D_x y,\ D_x f,\ \frac{df}{dx}.$$

The process of finding the derivative f' of a function f is called ***differentiation***.

2.2 DEFINITION. A function f is said to be ***differentiable at*** x_1 if $f'(x_1)$ exists. If f is differentiable at every point x_1 in its domain, we say that f is ***differentiable***.

Differentiability and continuity. Consider a function f defined by the equation (see Figure 7)

$$f(x) = \begin{cases} 3x - 2, & \text{if } x \leq 3, \\ 10 - x, & \text{if } x > 3. \end{cases}$$

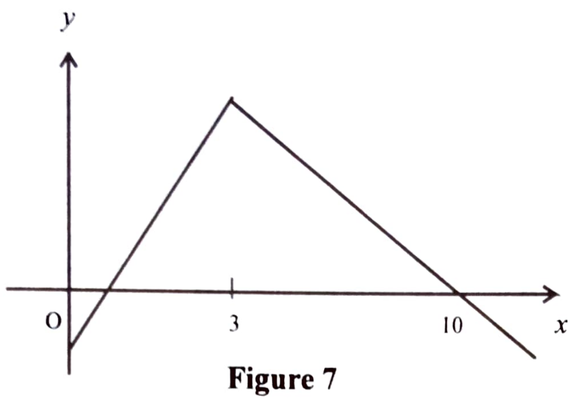

Figure 7

Since lim $f(x) = 7 = f(3)$, as $x \to 3$, it follows that f is continuous at $x = 3$. However, if we form the difference quotient,

$$\frac{f(3+\Delta x)-f(3)}{\Delta x} = \frac{f(3+\Delta x)-7}{\Delta x}$$

and calculate its limits as $\Delta x \to 0$ both from the left and from the right, we obtain,

$$\lim_{\Delta x \to 0^-} \frac{f(3+\Delta x)-7}{\Delta x} = 3, \text{ while } \lim_{\Delta x \to 0^+} \frac{f(3+\Delta x)-7}{\Delta x} = -1.$$

Since the left-hand and right-hand limits of the difference quotient are not equal, the limit of the difference quotient does not exist; that is, $f'(3)$ does not exist or f is not differentiable at $x = 3$. If we look at the graph of f in Figure 7, particularly at the point where $x = 3$, the nonexistence of the derivative is indicated by the absence of the tangent line there.

2.3 DEFINITION. The ***derivative from the right*** of a function f is

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

The ***derivative from the left*** of a function f is

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

Thus, for the function graphed in Figure 7, $f_-'(3) = 3$ and $f_+'(3) = -1$. Hence, $f'(3)$ cannot exist.

REMARK. The derivative $f'(x)$ exists and has the value A if and only if both of the one-sided derivatives exist and have the same value A.

2.3.1 EXAMPLE. Let the function f be defined by,

$$f(x) = \begin{cases} x^2, & \text{if } x \le 0, \\ -x^2, & \text{if } x > 0. \end{cases}$$

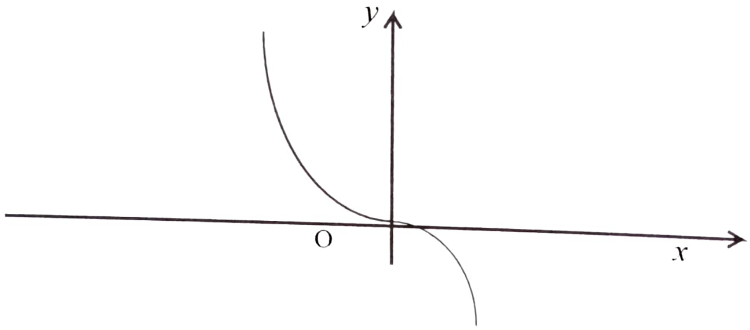

Figure 8

Determine if f is differentiable at $x = 0$ (Figure 8).

Solution. By Definition 2.3, we have

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

$$= \lim_{\Delta x \to 0^+} -\Delta x = 0.$$

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

$$= \lim_{\Delta x \to 0^-} \Delta x = 0. \quad \#$$

Since $f_+'(0) = f_-'(0) = 0$, $f'(0)$ exists and equals 0. Hence f is differentiable at $x = 0$. This example shows that a function defined "piecewise" can have a derivative at the boundary number between the "pieces".

Geometrically, a function that is differentiable at the point $(x_1, f(x_1))$ of its graph has a tangent line with slope $f'(x_1)$ at that point. Obviously, f is continuous there.

2.4 THEOREM. *If a function is differentiable at the number* x_1, *then it is continuous at* x_1.

Before we prove this, let us recall the definition of a continuous function at a point.

2.5 DEFINITION. The function f is said to be ***continuous at the number*** *a* if

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

Proof of Theorem 2.4. Assume that f is differentiable at x_1. We write,

$$f(x) - f(x_1) = \frac{f(x) - f(x_1)}{x - x_1}(x - x_1), \quad x \neq x_1.$$

We take the limits of both sides and apply some theorems on limits:

$$\lim_{x \to x_1}[f(x) - f(x_1)] = \lim_{x \to x_1} \frac{f(x) - f(x_1)}{x - x_1}(x - x_1)$$

$$= f'(x_1) \lim_{x \to x_1}(x - x_1) = 0.$$

Thus $\lim_{x \to x_1} f(x) = f(x_1)$ and f is continuous at x_1. #

REMARK. The expressions $f'(x_1) = \lim_{\Delta x \to 0} \frac{f(x_1 + \Delta x) - f(x_1)}{\Delta x}$ and $f'(x_1) = \lim_{x \to x_1} \frac{f(x) - f(x_1)}{x - x_1}$ are equivalent. To see this, let $x = x_1 + \Delta x$.

Then $\Delta x \to 0$ is equivalent to $x \to x_1$.

Formulas for derivatives. Up to now, we differentiated functions by direct use of the formal definition of the derivative as a limit of a difference quotient. Differentiation in this process can be tedious. Relief from this technique is given to us by applying basic theorems. These theorems are proved using the definition of the derivative.

2.6 THEOREM (CONSTANT RULE). *The derivative of a constant function is the zero function. In symbols, if $f(x) = c$ for all x, where c is a constant, then $f'(x) = 0$ for all x.*

2.6.1 EXAMPLE. Let $f(x) = 8 - \pi$ for all x. Then $f'(x) = 0$. #

2.7 THEOREM (IDENTITY RULE). *The derivative of the identity function is the constant function 1. In symbols, if $f(x) = x$ for all x, then $f'(x) = 1$ for all x.*

2.8 THEOREM (POWER RULE). *If n is an arbitrary constant, then*

$$D_x x^n = nx^{n-1} \text{ or } \frac{dx^n}{dx} = nx^{n-1}.$$

2.8.1 EXAMPLE. Given $h(x) = \sqrt{x}$, find $D_x h$.

Solution. By Theorem 2.8, $D_x(\sqrt{x}) = D_x(x^{\frac{1}{2}}) = \frac{1}{2}x^{-\frac{1}{2}}$. #

2.9 THEOREM (SUM RULE OR ADDITION RULE) *The derivative of a sum is the sum of the derivatives of the summands. In symbols, if f and g are differentiable functions of x, then*

$$D_x\,(f(x) + g(x)) = D_x\, f(x) + D_x\, g(x).$$

2.9.1 EXAMPLE. Find $D_x\left(\frac{1}{x} + 5\right)$.

Solution. $D_x\,(\frac{1}{x} + 5) = D_x\,(\frac{1}{x}) + D_x\,(5) = D_x\,(x^{-1}) + 0 = -x^{-2} = \frac{-1}{x^2}$. #

2.10 THEOREM (PRODUCT RULE OR MULTIPLICATION RULE). *The derivative of the product of two functions is the first function times the derivative of the second function plus the derivative of the first function times the second function. In symbols, if f and g are differentiable functions of x, then*

$$D_x\,[f(x)g(x)] = f(x)\;D_x\,[g(x)] + D_x\,[f(x)]\;g(x).$$

2.10.1 EXAMPLE. Find $D_x\,[(x+1)\,(x^2-2)]$

Solution. $D_x\,[(x+1)(x^2-2)] = (x+1)[\,D_x\,(x^2-2)] +$

$$+\,[\,D_x\,(x+1)](x^2-2)$$

$$= (x+1)\,(2x)\;+1(x^2-2)$$

$$= 2x^2+2x+x^2-2$$

$$= 3x^2+2x-2\quad\#$$

2.11 THEOREM (QUOTIENT RULE). *The derivative of a quotient of two functions is the denominator times the derivative of the numerator minus the numerator times the derivative of the denominator, all divided by the square of the denominator. In symbols, if f and g are differentiable functions of x, then*

$$D_x\left(\frac{f(x)}{g(x)}\right) = \frac{gD_xf(x) - fD_xg(x)}{g^2(x)} \quad\text{or}\quad \frac{d}{dx}\left(\frac{f(x)}{g(x)}\right) = \frac{g\frac{df}{dx} - f\frac{dg}{dx}}{g^2(x)}.$$

2.11.1 EXAMPLE. Find $D_x\left(\dfrac{x^2}{x^3+7}\right)$.

Solution. $D_x\left(\dfrac{x^2}{x^3+7}\right) = \dfrac{(x^3+7)[D_xx^2] - x^2[D_x(x^3+7)]}{(x^3+7)^2}$

$$= \frac{(x^3+7)(2x) - x^2(3x^2)}{(x^3+7)^2}$$

$$= \frac{14x - x^4}{(x^3 + 7)^2}. \quad \#$$

The next rule is a consequence of the product and the quotient rules.

2.12 THEOREM. *If c is any constant and f is a function of x then*

(*i*) $D_x\,(cf) = c\,D_x\,f$,

(*ii*) $D_x\,(\frac{c}{f}) = -\,c\,\frac{D_x f}{f^2}$.

2.12.1 EXAMPLE. Given $f(x) = 8x - 2/x$, find f'.

Solution. $D_x\,(8x - 2/x) = D_x\,(8x) - D_x\,(2/x)$

$$= 8\,D_x\,x - \frac{-2D_x x}{x^2} = 8 + 2/x^2. \quad \#$$

Derivative of a composite function. Suppose that $y = f(u)$ and $u = g(x)$. We can combine the two equations and write $y = h(x) = f(g(x))$. Here, the function h obtained by "chaining" f and g together defines a new function called a ***composition*** of f and g and is written $h = f \circ g$. For example, let

$y = u^2$ and $u = 5x + 1$.

Then, substituting the value of u from the second equation into the first equation, we get

$y = (5x + 1)^2$.

Thus, $y = h(x) = f(g(x)) = (f \circ g)(x) = (5x + 1)^2$. We now define composition of function precisely.

2.13 DEFINITION. Let f and g be two functions satisfying the condition that at least one number in the range of g belongs to the domain of f. Then the ***composition*** of f and g, in symbols $f \circ g$, is the function defined by the equation

$(f \circ g)(x) = f(g(x))$

2.13.1 EXAMPLE. Let $f(x) = 3x - 1$ and $g(x) = x^3$. Find,

(a) $(f \circ g)(2)$ (d) $(g \circ f)(x)$

(b) $(g \circ f)(2)$ (e) $(f \circ f)(x)$

(c) $(f \circ g)(x)$ (f) $(f \circ g)(3.007)$

Solution. (a) $(f \circ g)(2) = f(g(2)) = f(2^3) = f(8) = 23$;

(b) $(g \circ f)(2) = g(f(2)) = g(3(2) - 1) = g(5) = 125$;

(c) $(f \circ g)(x) = f(g(x)) = f(x^3) = 3x^3 - 1$;

(d) $(g \circ f)(x) = g(f(x)) = g(3x - 1) = (3x - 1)^3$;

(e) $(f \circ f)(x) = f(f(x)) = f(3x - 1) = 9x - 4$;

(f) $(f \circ g)(3.007) = f(g(3.007)) = f(3.007^3) \approx 80.57$.

An important result for finding the derivative of a composite function is the following rule known as the chain rule.

2.14 THEOREM (THE CHAIN RULE). *If $y = y(u)$ is a differentiable function of u and $u = u(x)$ is a differentiable function of x, then y is a differentiable function of x and*

$$D_x\, y = D_x y\ D_x u, \text{ or } \frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx}.$$

2.14.1 EXAMPLE. If $y = u^2$ and $u = 5x + 1$, find dy/dx.

Solution. $\dfrac{dy}{dx} = \dfrac{dy}{du}\dfrac{du}{dx} = (2u)(5) = 10u = 10(5x + 1)$. #

One way to understand the meaning of the chain rule is to think of it in terms of rates of change. The equation

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

just says that the rate of change of y per unit change in x is equal to the rate of change of y per unit change in u times the rate of change of u per unit change in x. For instance, if y is increasing twice as fast as u and u in turn is increasing thrice as fast as x, then y is increasing six times as fast as x.

In another notation, if we let $y = f(u)$, then $D_u f(u) = f'(u)$, and the chain rule takes the form,

$$D_x f(u) = f'(u)\, D_x\, u.$$

The above form is probably the most practical form of the chain rule for routine calculations of derivatives.

2.14.2 EXAMPLE. Find $D_x\,(x^3 - 6x)^2$ by using the chain rule.

Solution. Let $f(u) = u^2$ and $u = x^3 - 6x$ then $f(u) = (x^3 - 6x)^2$, and $f'(u) = 2u$, and $D_x\, u = D_x\,(x^3 - 6x) = 3x^2 - 6$. Hence,

$$D_x\,(x^3 - 6x)^2 = D_x f(u) = f'(u)\ D_x\, u$$

$$= 2u(3x^2 - 6) = 2(x^3 - 6x)(3x^2 - 6). \quad \#$$

To differentiate the expression $(x^3 - 6x)^2$ in another way, we can expand it first and then differentiate it.

One way of remembering the chain rule, given to us in [1], is stated as

$$Df(\text{whatever}) = f'(\text{whatever}) \times D(\text{whatever}).$$

For instance, using this device with f as the power function, we would have

$$D_x\,(\text{whatever})^2 = 2(\text{whatever})^{2\text{-}1}\, D_x\,(\text{whatever}),$$

Thus, the calculation of Example 3 could have been abbreviated as follows:

$$D_x\,(x^3 - 6x)^2 = 2(x^3 - 6x)^{2-1}\ D_x\,(x^3 - 6x) = 2(x^3 - 6x)\,(3x^2 - 6).$$

The chain rule is often used to calculate derivatives of the form $D_x u^n$, where $u = u(x)$ is a differentiable function of x and n is an arbitrary constant. Thus, letting $f(u) = u^n$, so that $f'(u) = nu^{n-1}$ we obtain the important formula

$$D_x u^n = nu^{n-1} D_x u.$$

In other words, we have the pattern

$$D(\text{whatever})^n = n(\text{whatever})^{n-1} D(\text{whatever}).$$

2.14.3 EXAMPLE. If $F(x) = \dfrac{1}{(4x-7)^3}$, find $F'(x)$.

Solution. $F'(x) = D_x \dfrac{1}{(4x-7)^3} = D_x (4x-7)^{-3}$

$= -3(4x-7)^{-4} D_x (4x-7) = -12(4x-7)^{-4}$. #

2.14.4 EXAMPLE. If $h(t) = \sin^4 t$, find $D_t h(t)$.

Solution. $D_t h(t) = D_t \sin^4 t = 4(\sin t)^3 D_t (\sin t) = 4(\sin t)^3 \cos t$.

2.14.5 EXAMPLE. Find $D_x G(x)$, if $G(x) = (x+3)(1-2x)^5$.

Solution. $G(x)$ is a product so we first use Theorem 2.10.

$$\begin{aligned}
D_x G(x) &= D_x [(x+3)(1-2x)^5] \\
&= (x+3) D_x (1-2x)^5 + [D_x (x+3)](1-2x)^5 \\
&= (x+3)5(1-2x)^4 D_x (1-2x) + (1-2x)^5 \\
&= (x+3)5(1-2x)^4 (-2) + (1-2x)^5 \\
&= (1-2x)^4((1-2x) - 10(x+3)) \\
&= (1-2x)^4(-29-12x). \quad \#
\end{aligned}$$

The next examples illustrate the repeated use of the chain rule.

2.14.6 EXAMPLE. Find $D_x\left(\sqrt{x^3-4}\right)^5$.

Solution. A first use of the chain rule gives,

$$D_x\left(\sqrt{x^3-4}\right)^5 = 5\left(\sqrt{x^3-4}\right)^4 D_x\left(\sqrt{x^3-4}\right)$$

$$= 5(x^3-4)^2 D_x\left(\sqrt{x^3-4}\right).$$

Applying the chain rule again on $D_x\left(\sqrt{x^3-4}\right) = D_x\left(x^3-4\right)^{\frac{1}{2}}$, we get,

$$D_x\left(\sqrt{x^3-4}\right)^5 = 5(x^3-4)^2 D_x\left(x^3-4\right)^{\frac{1}{2}}$$

$$= 5(x^3-4)^2 \left(\frac{1}{2}\right)\left(x^3-4\right)^{-\frac{1}{2}}(3x^2)$$

$$= \left(\frac{15x^2}{2}\right)\left(x^3-4\right)^{\frac{3}{2}}. \quad \#$$

If $u = u(x)$ is a differentiable function of x, we can combine the chain rule with the usual differentiation rules for trigonometric and exponential functions. Listed below are some of the most basic differentiation formulas:

$D_x \sin u = \cos u \; D_x u,$

$D_x \cos u = -\sin u \; D_x u,$

$D_x \tan u = \sec u \; D_x u,$

$D_x \sec u = \sec u \tan u \; D_x u,$

$D_x b^u = b^u \log b \; D_x u.$

3.14.7 EXAMPLE. Find the derivative of $f(x) = \tan(3x^6)$.

Solution. $f'(x) = \sec^2(3x^6)\; D_x\,(3x^6)$

$$= \sec^2(3x^6)\; 18x^5 = 18x^5 \sec^2(3x^6). \quad \#$$

2.14.8 EXAMPLE. If $F(\theta) = (1 - \tan(3\theta + \pi))^3$, find $D_\theta F$.

Solution. $D_\theta F(\theta) = D_\theta (1 - \tan(3\theta + \pi))^3$

$$= 3(1 - \tan(3\theta + \pi))^2 \, D_\theta (1 - \tan(3\theta + \pi))$$

$$= 3(1 - \tan(3\theta + \pi))^2 \, [-\sec^2(3\theta + \pi)] \, D_\theta (3\theta + \pi)$$

$$= 3(1 - \tan(3\theta + \pi))^2 \, [-\sec^2(3\theta + \pi)] \, (3)$$

$$= 9(1 - \tan(3\theta + \pi))^2 \, [-\sec^2(3\theta + \pi)]. \quad \#$$

As stated earlier, the chain rule is actually a rule for differentiating the composition of functions. To see this, let $y = f(u)$ and $u = g(x)$, so that $y = f(g(x)) = (f \circ g)(x)$. Hence, we have

$$(f \circ g)'(x) = \frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx} = f'(u)g'(x) = f'(g(x))g'(x).$$

REFERENCES

[1] Foulis, D. J. and Munem, M. A., *Calculus With Analytic Geometry*, 2e, Wadsworth, 1984.

[2] Callahan, J. et. al., *Calculus In Context*, Preliminary Edition, W. H. Freeman, 1993.

[3] Lee, P.Y., *Calculus*, Dragon Publishing, Singapore, 1993.

[4] Leithold, L., *The Calculus With Analytic Geometry*, 4e, Harper & Row, 1981.

[5] Logsdon, M., *A Mathematician Explains*, The University of Chicago Press, 1961.