13.19 THE BABYLONIAN ALGORITHM REVISITED

Terms, Symbols & Concepts Introduced
1. $\dfrac{d^2y}{dx^2}$
2. Chain rule for differentiation
3. Difference rule for differentiation of a polynomial
4. Newton-Raphson method
5. Second derivative
6. Sum rule for differentiation of a polynomial
7. Turning point

In this lesson, we will use what we have discovered about the Differential Calculus to re-investigate the Babylonian Algorithm that calculates both square and cube roots.

THE DERIVATIVES OF SIMPLE ALGEBRAIC FUNCTIONS

The derivative of a real-valued function determines precisely, i.e., with perfect accuracy, the behavior, or direction, of a continuous curve at every one of its infinitude of points. It is no wonder that mathematicians have applied the method of increments to virtually every kind of mathematical function. These derivatives, therefore, are an essential tool in the hands of physicists in their analysis and work with these functions.

Let's return to the function $y = 16x^2$ where, from the previous lesson, we showed that $y' = 32x$. This time, to complete the full picture of the parabola, we will consider negative values of x. By method of increments, we can also calculate the derivative of this function at several specific points. Verify the derivative answers in your notebook.

Investigate Table 1.

Table 1

x	-8	-4	-2	0	2	4	8
$y = f(x)$	1024	256	64	0	64	256	1024
$y' = f'(x)$	-256	-128	-64	0	64	128	256

This table tells us several facts about the curve. The y row reveals the y-coordinate for each x-coordinate, the image for each argument. We can plot these points. The y' row shows the slope of the line tangent to the curve at these images. By doing so, the derivative gives us information about (1) the direction the curve is going and (2) how steep it is.

- A slope of -256 means that the tangent line is pointing very steeply downward (\) at the coordinate (-8, 1024).
- A slope of -128 means that the tangent line is still pointing downward but not as steep at the coordinate (-4, 256).
- A slope of -64 means that the tangent line is still pointing downward but not as steep at the coordinate (-2, 64); i.e., the curve is rounding out as the argument approaches 0 from the left.
- A slope of 256 means that the tangent line is pointing very steeply upward (/) at the coordinate (8, 1024).
- A slope of 128 means that the tangent line is still pointing upward but not as steep at the coordinate (4, 256).
- A slope of 64 means that the tangent line is still pointing upward but not as steep at the coordinate (2, 64); i.e., the curve is rounding out as the argument approaches 0 from the right.

A slope of 0 means that the tangent line at (0, 0) is parallel (—) to the x-axis.[1] The tangent line, in this case, is the x-axis. The curve is turning directions at this point. This point is, therefore, a **turning point**.[2] Positive slopes mean that the tangent line is pointing upward (/). An analysis of the derivative tells us the shape of the parabola *without plotting its points*. With this example, by analyzing the derivative of a function, a mathematician or physicist can quantify the behavior of any curve at any point.

Fortunately, we do not have to apply the method of increments at every point to determine the derivative at that point. We can use the method of increments to derive general formulas that apply to different types of functions. Let's see how this works.

We already determined how to find the derivative of a linear function, i.e., $y = f(x) = ax + b \Rightarrow y' = a$.

Let's begin with a simple parabola, $y = f(x) = x^2$. Let's add Δx, a little bit of x, to x, which adds a little bit of y, Δy, to y. We get:

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

Applying the Binomial Theorem to the right side of the equation, we get:

$$y + \Delta y = (x + \Delta x)^2 \Leftrightarrow y + \Delta y = x^2 + 2x\Delta x + (\Delta x)^2$$

Since $y = x^2$, we can cancel these terms, i.e., subtract them, from both sides of the equation:

$$y + \Delta y = x^2 + 2x\Delta x + (\Delta x)^2 \Leftrightarrow \Delta y = 2x\Delta x + (\Delta x)^2$$

Next, we divide both sides of the equation by Δx to get the difference quotient:

$$\Delta y = 2x\Delta x + (\Delta x)^2 \Leftrightarrow \frac{\Delta y}{\Delta x} = 2x + \Delta x$$

Let's now let Δx get infinitesimally small. We get:

$$\frac{dy}{dx} = y' = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \left(2x + \Delta x\right) = \lim_{\Delta x \to 0} 2x + \lim_{\Delta x \to 0} \Delta x = 2x + 0 = 2x$$

We conclude: Given $y = f(x) = x^2$, then $y' = f'(x) = 2x$.

Let's now differentiate $y = f(x) = x^3$. First, add Δx to x:

$$y + \Delta y = (x + \Delta x)^3$$

Applying the Binomial Theorem/Pascal's Triangle to the right side of the equation (Review the relevant homework from Lesson 13.17.), we get:

$$y + \Delta y = (x + \Delta x)^3 \Leftrightarrow y + \Delta y = x^3 + 3x^2\Delta x + 3x(\Delta x)^2 + (\Delta x)^3$$

Since $y = x^3$, we can cancel these terms, i.e., subtract them, from both sides of the equation:

$$y + \Delta y = x^3 + 3x^2\Delta x + 3x(\Delta x)^2 + (\Delta x)^3 \Leftrightarrow \Delta y = 3x^2\Delta x + 3x(\Delta x)^2 + (\Delta x)^3$$

Next, we divide both sides of the equation by Δx to get the difference quotient:

---

[1] This observation turns out to be very important in solving maximum/minimum problems. See the relevant homework exercises in this lesson.

[2] The turning point is also known as the point of infection. In Latin, infection means a "slight dip."

$\Delta y = 3x^2\Delta x + 3x(\Delta x)^2 + (\Delta x)^3 \Leftrightarrow \dfrac{\Delta y}{\Delta x} = 3x^2 + 3x\left(\Delta x\right) + \left(\Delta x\right)^2$

Let's now let Δx get infinitesimally small. We get:

$$\frac{dy}{dx} = y' = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0}\left(3x^2 + 3x\Delta x + \left(\Delta x\right)^2\right) = \lim_{\Delta x \to 0} 3x^2 + \lim_{\Delta x \to 0} 3x\Delta x + \lim_{\Delta x \to 0}\left(\Delta x\right)^2 =$$
$$3x^2 + 0 + 0 = 3x^2$$

We conclude: Given $y = f(x) = x^3$, then $y' = f'(x) = 3x^2$.

We can apply the same process to $y = f(x) = x^4$. Try it on your own. The derivative y' should look like this:

$y' = f'(x) = 4x^3$

What pattern do we see?

Table 2	
Function	**$y' = f'(x)$**
$y = f'(x) = x^2$	$2x$
$y = f'(x) = x^3$	$3x^2$
$y = f'(x) = x^4$	$4x^3$

Do you think differentiating x^5 gives $5x^4$? Differentiating x^6 gives $6x^5$? You are right. In general:

$$\boxed{y = x^n \Rightarrow y' = f'(x) = nx^{n-1}}$$

What if n is negative or if n is a fraction? Try a few examples, e.g., $n = -2$ and $n = \frac{1}{2}$. The formula holds.[3] It is thrilling to see how working knowledge of the arithmetic of negative numbers and fractions bears such fruit in situations like these.

We have already noted that the derivative of a constant is 0; i.e., $y = f(x) = 5 \Rightarrow y' = f'(x) = 0$. The slope of the graph of $y = 5$ is a straight line parallel to the x-axis and intersects the y-axis at (0, 5). The slope of this line is 0. What about differentiating $y = f(x) = x^2 + 5$? Intuitively, we would think to sum the derivatives of the individual terms; i.e., $2x + 0 = 2$. Let's apply the method of increments to make sure. First add Δx to x:

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

Applying the Binomial Theorem to the right side of the equation, we get:

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

Since $y = x^2 + 5$, we can cancel these terms, i.e., subtract them, from both sides of the equation:

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

We proceed as before to calculate the derivative as $2x$.

Returning to our free-fall function, what happens in the case of $y = f(x) = 16x^2$? Our work will follow the algebraic derivation in the last lesson.

---

[3] If $n < 0$, we must assume $x \neq 0$. Can you explain why?

First, we add Δx to x:

$y + \Delta y = 16(x + \Delta x)^2$

Applying the Binomial Theorem to the right side of the equation, we get:

$y + \Delta y = 16(x + \Delta x)^2 \Leftrightarrow y + \Delta y = 16[x^2 + 2x\Delta x + (\Delta x)^2]$

Applying the Distributive Law to the right side of the equation, we get:

$y + \Delta y = 16[x^2 + 2x\Delta x + (\Delta x)^2] \Leftrightarrow y + \Delta y = 16x^2 + 32x\Delta x + 16(\Delta x)^2$

Since $y = 16x^2$, we can cancel these terms, i.e., subtract them, from both sides of the equation:

$y + \Delta y = 16x^2 + 32x\Delta x + 16(\Delta x)^2 \Leftrightarrow \Delta y = 32x\Delta x + 16(\Delta x)^2$

Next, we divide both sides of the equation by Δx to get the difference quotient:

$$\Delta y = 32x\Delta x + 16(\Delta x)^2 \Leftrightarrow \frac{\Delta y}{\Delta x} = 32x + 16(\Delta x)$$

Let's now let Δx get infinitesimally small. We get:

$$\frac{dy}{dx} = y' = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} (32x + 16\Delta x) = \lim_{\Delta x \to 0} 32x + \lim_{\Delta x \to 0} 16\Delta x = 32x + 0 = 32x$$

POSITION, VELOCITY, AND ACCELERATION FUNCTIONS

Therefore, $y = f(x) = 16x^2 \Rightarrow y' = f'(x) = 32x$. Here is our connection that we mentioned at the end of the last lesson. These equations match Galileo's observations. If x = time, and y = distance, then y', the derivative of distance concerning time, is the formula for velocity. That is, $v = y'$ where v = velocity. The derivative, in this case, represents the rate at which the distance, the position function, is changing in relation to time, exactly what velocity means.

Now let $v = 32x$. What is the derivative of v? That is, what is the rate at which the velocity is changing in relation to time? Since $v = 32x$ is a linear equation, then $v' = 32$. The constant 32 is the rate at which the velocity is increasing every second, the acceleration constant. If a = acceleration, then $a = v' = 32$. Galileo discovered this constant experimentally (See the relevant homework exercises in Lesson 12.12.). It represents the pull of the force of Earth's gravity upon a falling object. The Earth pulls an object down at a rate of 32 feet per second every second, or 9.8 meters per second every second. We note this relationship in Table 3, where $y = f(x) = 16x^2$ and we let x = 0, 1, 2, 3, 4, 5 seconds.

Table 3

x	**0**	**1**	**2**	**3**	**4**	**5**
y	0	16	64	144	256	400
v = y′	0	32	64	128	160	192
a = v′	0	32	32	32	32	32

Think about what this table reveals. The Differential Calculus accounts for what happens when we drop a stone from a building, a dramatic connection between mathematics and the physical world.

In summary, the position function representing the distance y a stone falls in a certain amount of time x is:

$y = f(x) = 16x^2$ [The rule is (1) square, then (2) multiply by 16.]

Differentiating y with respect to x gives us the velocity v, the rate at which the distance is changing in relation to time, which interpenetrates the slope of the tangent line to the graph of the function $y = f(x) = 16x^2$:

$v = y' = f'(x) = 32x$ (The rule is multiply by 32.)

Isn't this wonderful perichoresis? The derivative is the answer to two seemingly unrelated problems:

1. How do we draw a tangent line to a curve?
2. How do we compute the velocity of an object in motion?

Differentiating v with respect to x gives us the acceleration a, the rate at which the velocity is changing in relation to time:

$a = v' = y'' = f''(x) = 32$ (The rule is always 32.)

Note the y'' and $= f''(x)$ symbol. We say, "y double prime" and "f double prime of x." These symbols mean the second derivative. y', by implication, means the first derivative. Using the symbolism of Leibniz, we write the **second derivative** as follows:

$$a = \frac{d^2y}{dx^2}$$

The constant of acceleration, the second derivative of position, gives commentary to the way the universe of the Triune God's making coheres. The Father, Son, and Holy Spirit faithfully sustain the movement of bodies in the heavens and on the Earth in such a way that the second derivative of position unveils a constant that is an echo of the Triune God's covenant faithfulness in Christ the Son, the *logos*, who upholds all things (Hebrews 1:1-3; Colossians 1:15-17).

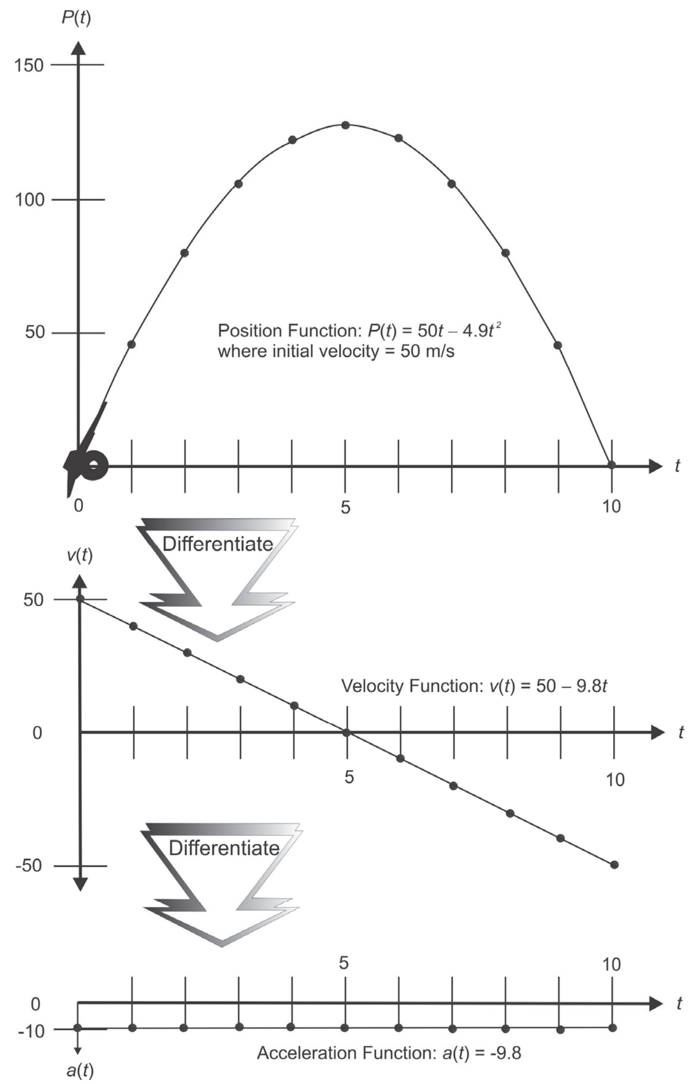

Figure 1. Physics equations for projectile motion where the pull of gravity, acceleration force in metric units, is a negative vector.

> In the context of the calculus, change in position is change that takes place along the spatial axis of a coordinate system, but the spatial axis itself may serve as a stand-in for *any change* that is made measurable by the real numbers, so that change in position functions as a large, a fabulously *general* concept, one standing in for change in *something*. It is the miracle of the calculus that change in something and change in time may be coordinated by means of the vastly greater abstraction of a function, purely an intellectual object, the key to the calculus, the key, in fact, to mathematics ….
>
> David Berlinski, *A Tour of the Calculus* (1995), pp. 63-64.

GENERAL DERIVATIVE FORMULA

Until you reach a constant, you can differentiate any function of the form ax^n as many times as you want. We can extend our derivation formula. In general:

$$\boxed{y = f(x) = ax^n \Rightarrow y' = f'(x) = nax^{n-1}}$$

With this formula in hand, we can calculate the first to the sixth derivatives of $y = f(x) = x^6$ as follows:

$y' = 6x^5$ (first derivative)

$y'' = \dfrac{d^2y}{dx^2} = 30x^4$ (second derivative)

$y''' = \dfrac{d^3y}{dx^3} = 120x^3$ (third derivative)

$y'''' = \dfrac{d^4y}{dx^4} = 360x^2$ (fourth derivative)

$y''''' = \dfrac{d^5y}{dx^5} = 720x$ (fifth derivative)

$y'''''' = \dfrac{d^6y}{dx^6} = 720$ (sixth derivative)

Although we compute these derivatives in mechanical fashion, we always need to remember that the n^{th} derivative function expresses the instantaneous rate of change of the $(n-1)^{\text{th}}$ function.

PERICHORESIS IN THE POLYNOMIAL WORLD

In the polynomial world (Lesson 12.11), when the derivative of a function is a constant, e.g., $y = f(x) = ax + b \Rightarrow y' = f'(x) = a$, the derivative ordinal tells us the degree of the function, the degree defined as the highest power of the polynomial's terms.[4] In our example, the first derivative is a constant. Therefore, the function is a linear equation of the first degree.

[4] We will do more work with polynomials in Lesson 14.3.

Inspect Table 4.

Table 4: General Polynomial Functions			
	Linear	**Quadratic**	**Cubic**
Form	$y = ax + b$	$y = ax^2 + bx + c$	$y = ax^3 + bx^2 + cx + d$
First Derivative	$y' = a$ (Constant)	$y' = 2ax + b$	$y' = 3ax^2 + 2bx + c$
Second Derivative		$y'' = 2a$ (Constant)	$y'' = 6ax + 2b$
Third Derivative			$y''' = 6a$ (Constant)
Degree of Polynomial	1	2	3

For a first degree polynomial, the first derivative is constant. For a second degree polynomial, the second derivative is constant. For a third degree polynomial, the third derivative is constant. In general, for an n^{th} degree polynomial, the n^{th} derivative is constant. We can also state that the derivative of a polynomial of degree $n > 0$ is a polynomial of degree $n - 1$. These observations are in perichoretic sync with our delta analysis work in Lesson 10.6, Lesson 11.6, and Lesson 11.11. Review that respective work now.

BACK TO BABYLON

We finally have the tools to use the Newton-Raphson method to finding the square root of a number. We have used many techniques to solve a quadratic equation of the form $ax^2 + bx + c = 0$. For example, we can use the inverse method, factoring, or completing the square/quadratic formula. A less complicated approach using Calculus leads us to a technique that solves not only quadratic equations, but even more complicated equations like cubics, quartics (fourth-degree polynomials), etc. All we need to do is write the equation in the form $f(x) = 0$ and find $f'(x) = 0$. This method interpenetrates the ancient Babylonian recursion formula for calculating square roots. In review, here are the steps to find $\sqrt{2}$ using this ancient method, where x_n and x_{n+1} are exchanged with I and O respectively:

Step 1. Let I = input, the initial guess of $\sqrt{2}$.

Step 2. Let O = the output where $O = \dfrac{I + \dfrac{2}{I}}{2}$.

(Note: We expect O to be a better approximation than I.)

Step 3. Recursive Step: Set the output equal to the input, $I = O$, and go to Step 2.

As we repeat this process, we approach the value of $\sqrt{2}$ as a limit. In theory, we can run this algorithm forever, but in practice, we stop it in a finite number of steps when we are satisfied with the degree of precision obtained.

DERIVATION OF THE NEWTON-RAPHSON METHOD

In functional terms, we set $f(x) = x^2 - 2$ and we seek a positive solution to x when $f(x) = 0$, i.e., we solve this equation for x:

$$x^2 - 2 = 0$$

Graphically (Figure 2), we seek to obtain the x-coordinate of the image where the curve intersects the x-axis. Given an initial guess I of this x-coordinate, i.e., x_1, the figure shows how to use the idea of a tangent line to generate a better guess O. Note how much closer O is to the desired solution than I. In other words, the graph tells us how to get the better estimate O, i.e., x_2, that is closer to the actual solution at x. Given I, we let O equal the x-intercept of the line tangent to the graph of the function at

$[I, f(I)]$. Since this tangent line is the line of the slope $f'(I)$ passing through $(x_1, y_1) = ([I, f(I)]$, the equation of this line is:

$$y - f(I) = f'(I)(x - I)$$

{Note: We have replaced (x_1, y_1) with $([I, f(I)]$.}

Why? Remember that the slope a is measured by the ratio of Δy over Δx, or:

$$a = \frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1}$$

Therefore, solving for $y_2 - y_1$, we get:

$$a = \frac{y_2 - y_1}{x_2 - x_1} \Leftrightarrow y_2 - y_1 = a(x_2 - x_1)$$

In our case, we know one point $(x_1, y_1) = ([I, f(I)]$ and the value of the derivative $f'(I)$ at I.

The better estimate O is the x-coordinate x_2 of $(x_2, 0)$, where the tangent line intersects the x-axis. (Please do not get the "zero" and "O" mixed up!) Thus, the x-coordinate of O must satisfy $y_2 - y_1 = a(x_2 - x_1)$. We substitute these values into this equation:

$x_1 = I$
$y_1 = f(I)$
$x_2 = O$
$y_2 = 0$
$a = f'(I)$

From $y_2 - y_1 = a(x_2 - x_1)$, we get:

$$0 - f(I) = f'(I)(O - I)$$

Our object is to find O if we are given I. We apply these algebraic operations:

$$0 - f(I) = f'(I)(O - I) \Leftrightarrow$$
$$-f(I) = f'(I)(O - I)$$

Assuming $f'(I) \neq 0$, we divide both sides of the equation by $f'(I)$ and get:

$$-f(I) = f'(I)(O - I) \Leftrightarrow -\frac{f(I)}{f'(I)} = O - I$$

Solving for O, we get:

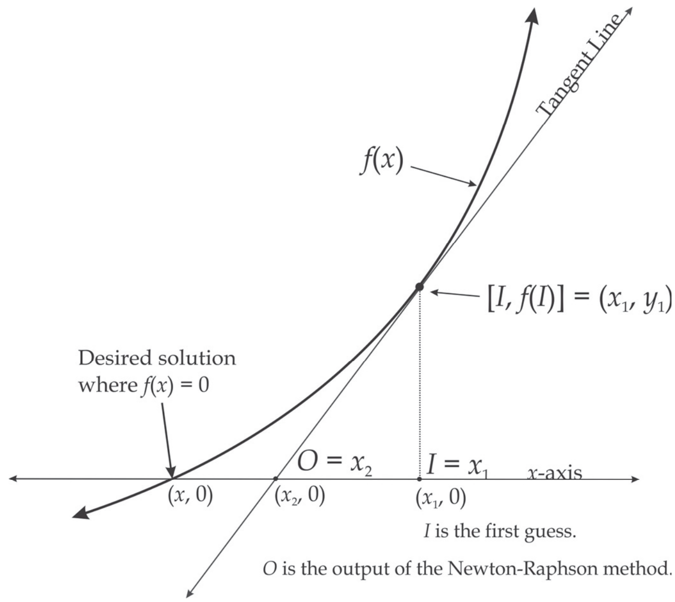

Figure 2. First iteration of the Newton-Raphson method.

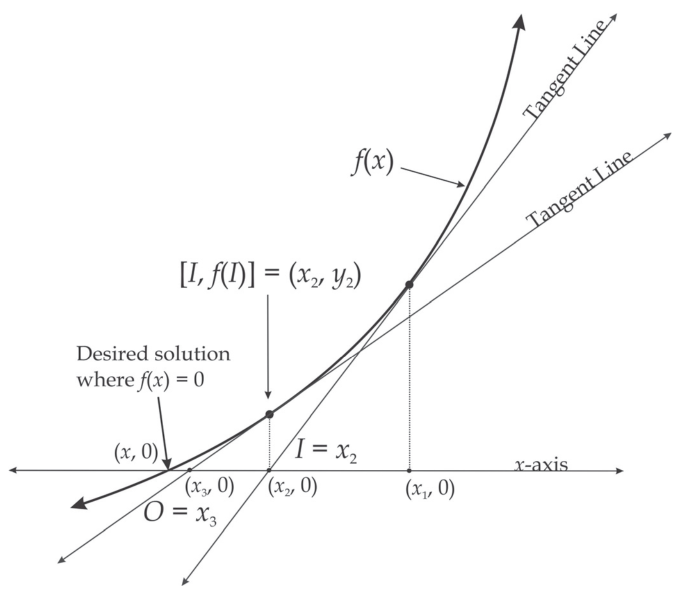

Figure 3. Second iteration of the Newton-Raphson method.

$$-\frac{f(I)}{f'(I)} = O - I \Leftrightarrow I - \frac{f(I)}{f'(I)} = O \Leftrightarrow$$

$$O = I - \frac{f(I)}{f'(I)}$$

The last equation expresses in Calculus terms the **Newton-Raphson method**, a recursive algorithm, of computing a better guess O for a given guess I as a solution to the equation $f(x) = 0$. We apply the Newton-Raphson method to solve an equation $f(x) = 0$ is as follows:

Step 1. Let I = the initial guess of the solution.

Step 2. Let O = output where $O = I - \frac{f(I)}{f'(I)}$.

(Note: We expect O to be a better approximation than I.)

Step 3. Recursive Step: Set the output equal to the input, $I = O$, and go to Step 2.

INSIGHT AND WARNING

We can apply the Newton-Raphson method, in principle, to any function $f(x) = 0$ if we know $f'(x)$. In practice, it is wise to sketch the curve first to make sure that it intersects the x-axis; i.e., it has a real number solution. If the curve does not intersect the x-axis, the method will give you garbage. Also, it is advisable to choose your initial guess I so that $f(I)$ is already close to 0. When you do this, the method will work with astonishing speed.

THE SQUARE ROOT OF 2

Let's apply the Newton-Raphson method to find $\sqrt{2}$, i.e., the positive square root of 2. First, we let $x = \sqrt{2}$. Squaring both sides, we get:

$$x = \sqrt{2} \Rightarrow x^2 = 2$$

Subtracting 2 from both sides, we get:

$$x^2 = 2 \Leftrightarrow x^2 - 2 = 0$$

We set $f(x) = x^2 - 2 \Rightarrow f'(x) = 2x$

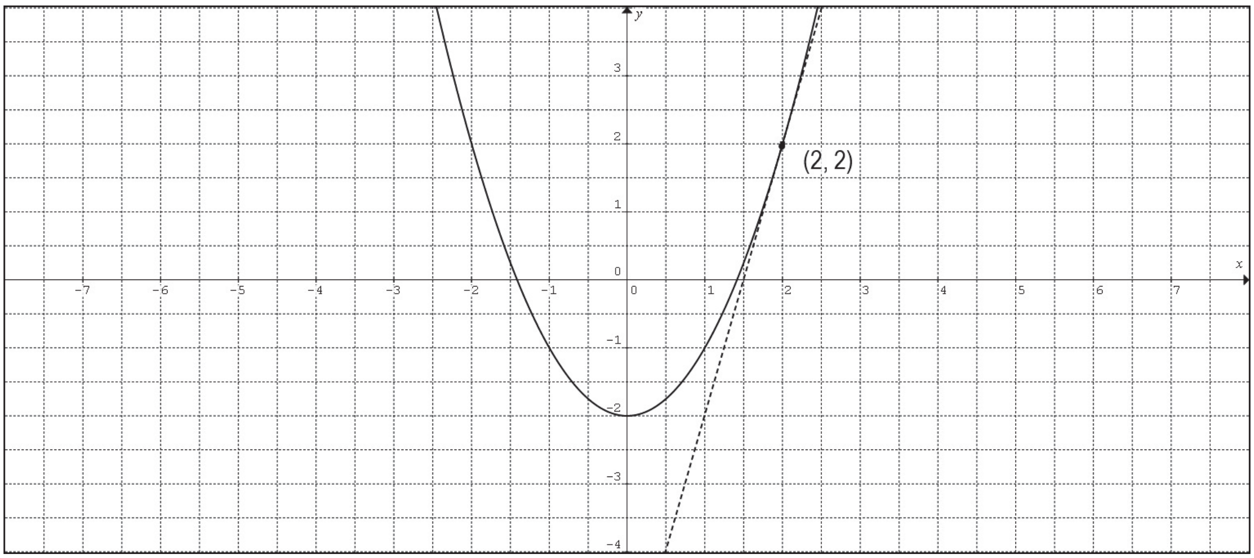

Figure 4. $y = x^2 - 2$

Since $f(x) = 0$, we invoke the Newton-Raphson method. First, we graph $y = x^2 - 2$. Good graphics software will plot the tangent line to any curve at a specific point. In Figure 4, we have shown the tangent line to the graph at (2, 2). We know by inspection that the positive solution is between 1 and 2. We choose 2 as our first guess.

Substituting the appropriate values into the Newton-Raphson algorithm, we do our first iteration:

$$O = I - \frac{f(I)}{f'(I)},\ I = 2, f(2) = 2, \text{ and } f'(2) = 4 \Rightarrow O = 2 - \frac{2}{4} = 2 - \frac{1}{2} = \frac{4-1}{2} = \frac{3}{2} = 1.5$$

Using the Babylonian algorithm, we get:

$$O = \frac{I + \frac{2}{I}}{2} \text{ and } I = 2 \Rightarrow O = \frac{2 + \frac{2}{2}}{2} = \frac{2+1}{2} = \frac{3}{2} = 1.5$$

The same result!

Now, for our second iteration. First, the Newton-Raphson algorithm:

$$O = I - \frac{f(I)}{f'(I)} \text{ and } I = 1.5, f(1.5) = 0.25, \text{ and } f'(1.5) = 3 \Rightarrow$$

$$O = 1.5 - \frac{0.25}{3} = 1.5 - 0.08\overline{3} = 1.41\overline{6}$$

Now, the Babylonian algorithm:

$$O = \frac{I + \frac{2}{I}}{2} \text{ and } I = 1.5 \Rightarrow O = \frac{1.5 + \frac{2}{1.5}}{2} = \frac{1.5 + 1.\overline{3}}{2} = \frac{2.8\overline{3}}{2} = 1.41\overline{6}$$

Again, the same result!

We continue with our third iteration. First, the Newton-Raphson algorithm:

$$O = I - \frac{f(I)}{f'(I)} \text{ and } I = 1.41\overline{6}, f(1.41\overline{6}) = 0.0069\overline{4}, \text{ and } f'(1.41\overline{6}) = 2.8\overline{3} \Rightarrow$$

$$O = 1.41\overline{6} - \frac{0.0069\overline{4}}{2.8\overline{3}} \approx 1.4142$$

Now, the Babylonian algorithm:

$$O = \frac{I + \frac{2}{I}}{2} \text{ and } I = 1.41\overline{6} \Rightarrow O = \frac{1.41\overline{6} + \frac{2}{1.41\overline{6}}}{2} \approx 1.4142$$

PERICHORESIS IN RECURSION

Both algorithms generate a very accurate and identical approximation of $\sqrt{2}$, iteration by iteration. The two algorithms, $O = \frac{I + \frac{2}{I}}{2}$ and $O = I - \frac{f(I)}{f'(I)}$ interpenetrate in stunning perichoresis! Why? Let's do more Algebra to prove that the two algorithms act in perichoresis because they represent the same expression.

First, let's simply the Babylonian algorithm:

$$O = \frac{I + \frac{2}{I}}{2} \Leftrightarrow O = \frac{\frac{I^2 + 2}{I}}{2} \Leftrightarrow O = \frac{I^2 + 2}{I} \div 2 \Leftrightarrow O = \frac{I^2 + 2}{2I}$$

Now, let's work the Newton-Raphson algorithm to see if we can get the same result. We do this:

$$O = I - \frac{f(I)}{f'(I)}, f(I) = I^2 - 2, \text{ and } f'(I) = 2I \Rightarrow O = I - \frac{I^2 - 2}{2I} \Leftrightarrow O = \frac{2I^2 - \left(I^2 - 2\right)}{2I} \Leftrightarrow$$

$$O = \frac{2I^2 - I^2 + 2}{2I} \Leftrightarrow O = \frac{I^2 + 2}{2I}$$

QED

Applying algebraic operations to both algorithms generates $O = \frac{I^2 + 2}{2I}$ and, therefore, our algebraic work reveals the reason for the perichoresis; i.e., $O = \frac{I^2 + 2}{2I}$ is contained in both $O = \frac{I + \frac{2}{I}}{2}$ and $O = I - \frac{f(I)}{f'(I)}$. $O = \frac{I + \frac{2}{I}}{2}$ and $O = I - \frac{f(I)}{f'(I)}$ interpenetrate each other in the form $O = \frac{I^2 + 2}{2I}$.

Pause to wonder!

CONCLUSION

In the last three lessons, we toured the landscape of Differential Calculus. Many university Calculus textbooks run over 1000 pages. In these lessons, we have given you a basic introduction to the way mathematicians reason about the infinitesimal. You have learned the elementary procedures of driving the Calculus car by traveling some backcountry roads. University Calculus is like driving on a freeway during rush hour traffic. In these courses, you will learn how to find derivatives of logs, exponential functions, and trigonometric functions. You will also learn a good number of differential techniques. Then, you will explore the inverse of differentiation, i.e., integration, and much, much more.

All this work is preparatory ground for the exploring more advanced methods like differential equations, i.e., equations with derivatives as terms (e.g., $\frac{dy}{dx} = 5x + 3$). From these equations, you will learn many techniques that open the workings of the physical world in manifold ways. For example, you will learn how Isaac Newton, in the 17^{th} century, used differential equations to pinpoint the velocity needed for a ballistic-type rocket to escape the gravitational pull of the Earth, well before the invention of such rockets in the 20^{th} century. The initial velocity applied to such rockets (e.g., in the NASA moon program in the 1960s and 1970s) confirmed that these objects do, indeed, escape the gravitational pull of the Earth according to Newton's mathematical analysis and predictions.

From differential equations, we enter the world of partial derivatives of surfaces. In the 19^{th} century, the Scottish scientist James Clerk Maxwell (1831-1879) used this type of mathematics to derive a series of beautiful and symmetric equations, equations that describe the laws of electricity and magnetism. He used these equations to predict the existence of electromagnetic waves and the speed of light. In 1887, the German physicist Heinrich Hertz (1857-1894) confirmed Maxwell's equations by generating and receiving electromagnetic waves in his laboratory. In 1901, the Italian inventor Guglielmo Marconi (1874-1937) transmitted radio waves across the Atlantic Ocean using the telegraph. Also, around this time, in the early 20^{th} century, Albert Einstein (1879-1955) employed Maxwell's work to develop his Special and General Theories of Relativity.

Electromagnetism is a fundamental force in the universe. It is the underlying reason for the way things work in the micro- and macro-realms. It is essential in technology, from microwave ovens to electronic watches to personal computers to smart phones to the World Wide Web. We can even explain our bodies, from electrochemical nerve impulses to the electric signals controlling our heartbeat, using Maxwell's principles.

Such wonders await you if you accept the challenge to continue your mathematical studies. I hope you do! One of you might be the next Maxwell!

> … this vast branch of mathematics, partial differential equations serves the purpose of expressing the basic physical principles of such prevalent and vital phenomena as sound, heat, the various forms of electromagnetic waves, water waves, vibrations in rods, the flow of fluids and gases, and so forth, and that from these differential equations we can deduce by mathematical methods alone a vast amount of information about these phenomena. In fact, it is fair to say that the subject of differential equations is today the heart of mathematics and it is certainly the most useful branch for the study of the physical world.
>
> Morris Kline, *Mathematics and the Physical World* ([1959] 1981), p. 422.

EXERCISES

Define the following words:

1. Turning point
2. Second derivative
3. Newton-Raphson method
4. Sum rule for differentiation of a polynomial
5. Difference rule for differentiation of a polynomial
6. Chain rule for differentiation of a polynomial

Prove the following by the method of increments unless stated otherwise:

7. $y = bx \Rightarrow \dfrac{dy}{dx} = b$.
8. $y = c$ (a constant) $\Rightarrow \dfrac{dy}{dx} = 0$.
9. $y = ax^3 \Rightarrow \dfrac{dy}{dx} = 3ax^2$.
10. $y = ax^2 + bx \Rightarrow \dfrac{dy}{dx} = 2ax + b$ (**Sum rule for differentiation**.)
11. $y = ax^2 - bx \Rightarrow \dfrac{dy}{dx} = 2ax - b$ (**Difference rule for differentiation**.)
12. $y = x^n \Rightarrow \dfrac{dy}{dx} = nx^{n-1}$ (Use Mathematical Induction.)

Answer these questions:

13. Why must x be non-zero when $n < 0$ in Question 12?
14. Find the derivative, the instantaneous rate of change of the dependent variable compared to the independent variable, of:
    (a) $y = \sqrt{z}$
    (b) $y = \sqrt[3]{v}$
    (Hint: Write your answer using exponents that are positive. Convert $\sqrt{z}$ and $\sqrt[3]{v}$ to fractional exponents first.)
15. (a) Apply the method of increments to find the instantaneous rate of change of $y = x^2 + 7$ and compare the result with the instantaneous rate of change of $y = x^2$.
    (b) What general conclusion do your results suggest?
16. Explain geometrically why the functions $y = x^2$ and $y = x^2 + 7$ should have the same derivative at, say, $x = 2$.
17. What is the slope of the line tangent to the following curves at $x = -1$:
    (a) $2x^2 - 3y = 8$
    (b) $2x^3 - 3y = 8$
    (c) $y = 3x^3 - 2x^2 - 6x + 2$
    (d) Graph all three functions together along with $x = -1$. Justify the slopes you calculated based on what you see on the graphs.

Find the derivative for the following functions:

18. $y = 3x^2$

> One cannot escape the feeling that these mathematical formulas [Maxwell's equations – JN] have an independent existence and an intelligence of their own, that they are wiser than we are, wiser even than their discoverers, that we get more out of them than was originally put into them.
>
> Heinrich Hertz, cited in Eric Temple Bell, *Men of Mathematics* ([1937, 1965] 1986), p. 16.

19. $y = (2/3)x^4$
20. $d = 2t^2$
21. $y = \dfrac{1}{2}x^2$
22. $k = 0.5p^4$
23. $y = 4x^3$
24. $y = -3x^2$
25. $d = -16t^2$
26. $y = x^{\frac{5}{2}}$
27. $y = x^{\frac{3}{2}} + 5x + 6$
28. $h = -4.9t^2 + 39.2t$
29. $h = 128t - 16t^2$

Write your answer using exponents that are positive when you find the derivative of:

30. $y = x^{-1}$
31. $k = 2t^{-2}$
32. $m = -6d^{-5}$
33. $y = 1.5x^{-(2/3)}$

Calculate the second derivative $\left(\dfrac{d^2y}{dx^2}\right)$ for the following functions:

34. $y = 4x^3$
35. $y = 128x - 16x^2$
36. $y = 8x^2$
37. $y = 32x$
38. $y = -19x^5$
39. $y = -2x^{-2}$ (State with positive exponent.)
40. $y = 3x^{-(4/5)}$ (State with positive exponent.)

Answer these questions:

41. The instantaneous acceleration of a moving object is by definition the instantaneous rate of change of the velocity with respect to the time. Suppose an object falls a distance d in t seconds governed by the formula $d = -4.9t^2$, where d is in meters. (Note: We are taking into account the downward motion of the object by the negative sign.) What is the instantaneous acceleration of the object at any time t?
42. If an object is thrown upward with the initial velocity of 100 ft/s, then the distance d it falls in t seconds is given by the formula $d = 100t - 16t^2$. Calculate the velocity of the object at the end of the fourth second of fall.
43. Suppose that an uphill path can be represented by the equation $y = \left(\dfrac{1}{100}\right)x^2$.
    (a) What is the slope of the hill at $x = 3$?
    (b) Is the slope significantly steeper or more gradual at $x = 3$ or $x = 5$?
    (c) Determine the slope at $x = 0$ and interpret the result geometrically.
44. Suppose that the path of a projectile is represented by the quadratic equation $y = 4x - x^2$.
    (a) What direction does the projectile have when $x = 1$?
    (b) What direction does the projectile have when $x = 3$?

(c) At what value of x is the direction of the projectile horizontal; i.e., slope = 0?

45. Figure 5 illustrates the variation of a certain function. Describe how the derivative of y with respect to x varies as x increases from A to B.

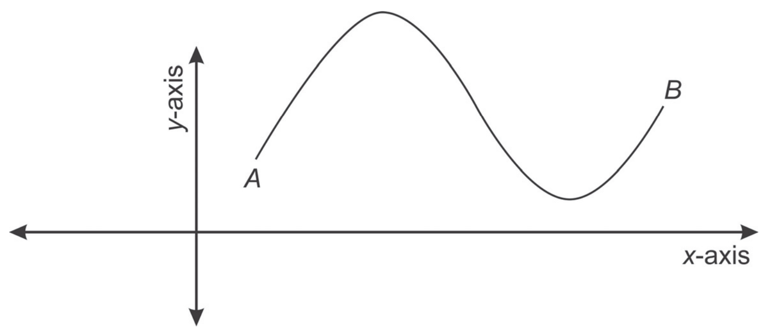

Figure 5.

How would you differentiate this function, $y = \sqrt{x^2 + a^2}$, where a is a constant? We can rewrite this function as follows: $y = \left(x^2 + a^2\right)^{\frac{1}{2}}$. This function is a three-stage composite function (Lesson 11.13) and the rule is (1) square, (2) add a^2, then (3) extract the square root of the sum. We can consider this three-stage function in two parts, (1) the two-stage function in the radicand $x^2 + a^2$ and (2) raising to the power of 1/2. Instead of trying to differentiate this function *in toto*, mathematicians attack it in two parts.

If we let $u = x^2 + a^2$, the problem becomes $y = u^{\frac{1}{2}}$. Differentiating y with respect to u gives us $y' = \frac{dy}{du} = \frac{1}{2}u^{-\frac{1}{2}}$. Next, differentiating u with respect to x gives us $u' = \frac{du}{dx} = 2x$. Then, $\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$. The derivatives appear to act like fractions, and they are in that they represent ratios, where we can cancel du from the equation: $\frac{dy}{dx} = \frac{dy}{\cancel{du}_{1}} \cdot \frac{\cancel{du}^{1}}{dx} = \frac{dy}{dx}$

This process is the **chain rule for differentiation**, a very useful short cut when differentiating complex functions. Applying the chain rule to our composite function, we get:

$$\frac{dy}{dx} = \frac{1}{2}u^{-\frac{1}{2}}2x = \frac{1}{2}\left(x^2 + a^2\right)^{-\frac{1}{2}}2x = x\left(x^2 + a^2\right)^{-\frac{1}{2}} = \frac{x}{\sqrt{x^2 + a^2}}$$

(Remember: $x^{-\frac{1}{2}} = \frac{1}{x^{\frac{1}{2}}} = \frac{1}{\sqrt{x}}$).

Find the first derivative of the following functions, where a is a constant, using the chain rule: (Note: These problems will give you a flavor of the intense algebraic work involved in Calculus.)

46. $y = \left(x^2 + a^2\right)^{\frac{3}{2}}$

47. $y = \left(2x^2 + a^2\right)^{\frac{5}{2}}$

48. $y = \sqrt{x-a}$

49. $y = \dfrac{1}{\sqrt{a+x^2}}$

50. $y = \dfrac{1}{\sqrt{x^3-a^2}}$

51. $y = \left(x + \sqrt{x^2+x+a}\right)$

52. Here is a superficial proof of the chain rule in terms of demonstrating how it works. Let's say, we have three children, Joy, Susan, and Charlotte. Joy grows twice as fast as Susan and Susan grows three times as fast as Charlotte. How much faster is Joy growing than Charlotte?

Prove the following:

53. Show how the Babylonian algorithm $O = \dfrac{2I + \dfrac{5}{I^2}}{3}$ for finding the cube root of 5 interpenetrates the Newton-Raphson method $O = I - \dfrac{f(I)}{f'(I)}$ for finding the cube root of 5.

Create Microsoft spreadsheet programs that apply the Newton-Raphson method to find the positive real number solution, unless otherwise indicated, to the following when $f(x) = 0$: (Write your answer to fifteen decimal places.)

54. $f(x) = x^2 - 3$
55. $f(x) = x^2 - 5$
56. $f(x) = x^2 - 92$
57. $f(x) = x^3 - 2$
58. $f(x) = x^3 - 3$
59. $f(x) = x^3 - 5$
60. $f(x) = x^3 - 24$
61. (a) $f(x) = -16x^2 - 50x + 200$ (Find all real number solutions.)
    (b) Graph to verify your results.

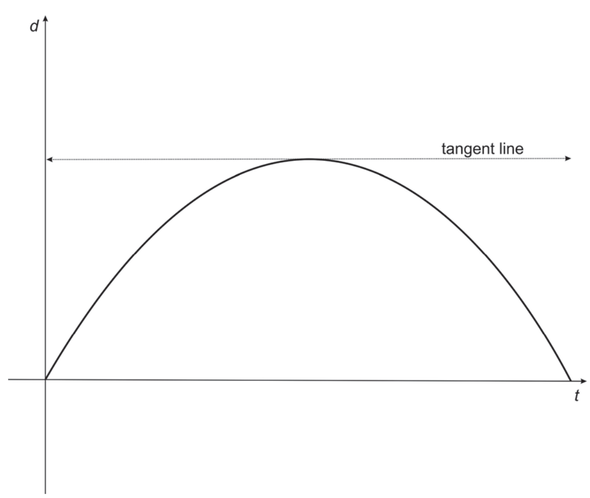

Figure 6.

Answer the following questions:

62. I pick up the stone and drop it into a "stone" cannon. I point the cannon straight up, ignite the gunpowder, and pull the trigger. Let's say that the cannon thrusts the stone upward with an initial velocity of 96 ft/s. Figure 6 is a general graph of the shape of a position function.
    (a) Write the position function *d* in terms of time *t*, taking in account the downward pull of gravity on the stone.
    (b) Inspect the figure and determine a method, using differential calculus, for finding the coordinates of the maximum height of the stone; i.e., the time *t* at which the distance *d* is the highest.
    (c) Using the method in (b) to find (t, d) where *d* is maximum.
    (d) Graph to verify your results.
63. Mr. City Slicker wants to build an ancient Egyptian amusement park on some property next to the Mr. Farmer's land. Mr. City Slicker is entertaining the nifty idea of transporting the park visitors around on camels. In fact, he has bought a herd of camels and wants to fence them in near a river that runs through his property. Mr. City Slicker can only afford 800 feet of fencing for his camel herd and he wants to enclose them in a rectangular field one side of which is bounded by the river. He finds a straight section of the river and figures that his idea is great because he doesn't have to worry about watering the camels. His problem is that he doesn't know how to get the most area out

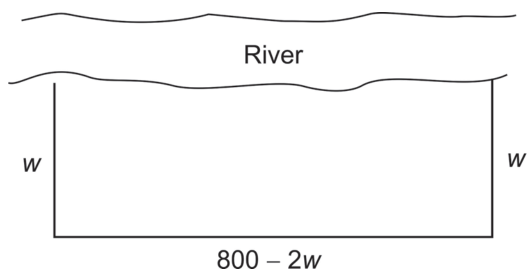

Figure 7.

of his 800 feet of fencing. He needs to know what to do before he starts to dig holes for his fence posts. "Mr. Farmer surely would know how to do this since he had such a nifty way of counting the number of cows in his pasture," thinks Mr. City Slicker to himself as he approaches Mr. Farmer's house and knocks on the screen door (Lesson 13.13). Mr. Farmer listens to Mr. City Slicker explain the situation. At the end of the account, Mr. Farmer pulls on his graying beard a few times in deep contemplative thought. Suddenly he exclaims, "Derivative!" He quickly retreats to his study. In the meantime, Mr. City Slicker stands in dumbfounded silence thinking, "Did I hear him right? I'm sure that he said "drive a Tiv." But I drive a Tercel! Is there a new model car that's out? And, how can driving a 'Tiv' solve my problem? I guess I'll have to trust Mr. Farmer because he sure knows how to count cows!" Meanwhile, back in the study, Mr. Farmer is drawing a diagram. Let's peak over his shoulder to find out what he is doing (Figure 7). Fortunately for us, Mr. Farmer always thinks out loud when he works in his study so we get to hear his reasoning. "Let's see," says Mr. Farmer, "I've got 800 feet of fence and three sides to work with. I've got to make a rectangular field for the camels so I'll let the width of the field be w. That takes care of two of the three sides and I've used up $2w$ feet of fence. I have $800 - 2w$ feet left. That's for the third side. The area of a rectangle is its length multiplied by its width, so I've got this equation:

$$A = w(800 - 2w) = 800w - 2w^2$$

(a) Finish Mr. Farmer's reasoning by finding w that gives the maximum area A.
(b) Graph to verify your results.

64. A farmer wishes to use 100 m of fencing to enclose a rectangular area and to divide the area into two rectangles by running a fence down the middle. (Hint: Draw a picture to represent the situation and then write down all relevant formulas.)
(a) What dimensions should he choose to enclose the maximum total area?
(b) Graph to verify your results.
65. We want to fence a small rectangular pen containing 24 square yards. The front, to be made of stone, will cost \$10 per yard of fencing, while each of the other three wooden sides will cost only \$4 per yard. What is the least amount of money that will pay for the fencing?
(a) Let L = the length of the front and C = cost. Determine a formula for C is terms of L.
(b) Find the minimal cost rounded to the nearest penny. (Note: Round L to the nearest hundredth.)
(c) Graph to verify your results.
66. In the previous question, C (total cost) can be expressed either in terms of L (length) alone or W (width) alone while L and W are connected in that $LW = 24$.
(a) Write again the formula for C in terms of L.
(b) Write a formula for L in terms of W.
(c) By substitution, write a formula for C in terms of W.
(d) Notice how the equations in (a) and (b) are links in a chain producing to the equation in (c). In other words, the equation is (a) is C as a function of L, i.e., $C(L)$, the equation in (b) is L as a function of W, i.e., $L(W)$, and the equation in (c) is C as a function of W, i.e., $C(W)$. Note the chain: $C \to L$, then $L \to W$, then $C \to W$. This chain illustrates, pun intended, the chain rule for differentiation. Calculate the three derivatives: $\frac{dC}{dL}, \frac{dL}{dW}$, and $\frac{dC}{dW}$.
(e) Apply the chain rule and show that $\frac{dC}{dL} \cdot \frac{dL}{dW} = \frac{dC}{dW}$. Isn't this a delightful connection?
67. To illustrate the power and efficiency of Calculus, we are going to prove that to maximize the area of a rectangle of perimeter p, then it must be a square.
(a) Let x and y be the dimensions of any rectangle, then write a formula for p in terms of x and y.
(b) Solve this formula for y.
(c) The area of any rectangle is given by $A = xy$. Express A as a function of x only and apply Calculus to find the maximum area. We can prove this relationship using Euclidean geometry, but this method is much more direct.
68. Does the function $y = x^3$ have a maximum or minimum value at $x = 0$? Plot the function using graphing software to assist you with your answer.

HOW A LITTLE MATHEMATICS BOOK INTERSECTED MY LIFE

Hungarian born Rózsa Péter (1905-1977) grew up in a land torn by war and civil strife. What we take for granted in terms of everyday living was for her *never easy*. Péter enrolled in Eötvös Loránd University in 1922 with the goal to earn a degree in chemistry. It did not take her long to discover the enchantment that mathematics offered. After changing majors, she studied under many professors who were world-famous mathematicians. She received her undergraduate degree in 1927, continued graduate studies, and earned her living by tutoring and teaching high school. She received her Ph. D. degree summa cum laude in 1935 and her studies pioneered the development of a new field of mathematics named recursive functions.

Figure 8. Rózsa Péter. Source: Wikimedia Commons.

World War II began in Europe in 1939. The Nazi juggernaut soon occupied Hungary and fascist laws forbade Péter to teach. She was even briefly confined in the Budapest ghetto. Amid suffering many hardships, including the painful loss of her brother along with many friends, students, and fellow mathematicians, she continued studying and writing. In the autumn of 1943, she finished a short mathematics manuscript. It was really a collection of letters she wrote to a friend explaining the nature of mathematics. Since no books could appear during the censorship imposed by the Nazis and since allied bombing destroyed many copies of this manuscript, the few copies that remained first appeared in 1945, on the first free book day.

Shortly after the end of the war, Budapest Teachers' College hired her. In 1951, she published an award-winning monograph on recursive functions. When the college closed in 1955, she became a mathematics professor at Eötvös Loránd University holding this post until her retirement in 1975. She would often speak on mathematics to general audiences. She entitled her lectures "Mathematics is Beautiful" saying, "No other field can offer, to such an extent as mathematics, the joy of discovery, which is perhaps the greatest human joy."[5]

In 1976 she published a book entitled *Recursive Functions in Computer Theory*. She died on the eve of her birthday in 1977. In her eulogy, one of her students recalled that she taught "that facts are only good for bursting open the wrappings of the mind and spirit" in the "endless search for truth."

What happened to that little manuscript that Péter wrote during the war years? It was finally published in Hungarian in 1957. The English translation was made by Dr. Z. P. Dienes and published in England by G. Bell and Sons, Ltd., in 1961. Simon and Schuster published it on the other side of the Atlantic in 1962. Dover Publications of New York, the great reprinter of out of print books, produced a wonderful paperback version in 1976 entitled *Playing with Infinity: Mathematical Explorations and Excursions*.

[5] For a transcript of one of these lectures, see Rózsa Péter, "Mathematics is Beautiful," *The Mathematical Intelligencer,* 12 (1990): 58-64.

I first encountered[6] the Dover version of this book in 1984 while teaching high school mathematics in the little Australian hamlet of Booleroo Centre. A small farming town with a population of 300 but serving a district population of nearly 2000, it still had hitching posts in front of local stores, individual stores where you could buy meat from a butcher, bread from a baker, everything else from Prests, a general merchandiser and grocer, and get a great meal at the local pub attached to the Booleroo Centre Hotel. It is in the state of South Australia, located in the southern Flinders Ranges, and about a hundred or so miles south of the start of the famous Australian outback territory.

> … Mathematics is an organic whole: wherever we touch it, connecting links from all other branches come crowding into our minds.
> Rózsa Péter, *Playing with Infinity* ([1957, 1961] 1976), p. 41.

Back to Péter's little book … At once I recognized her rare ability not only to explain mathematical topics but also to recognize and organize its intricate structure. Compared to the mathematical drivel taught by much of high school mathematics textbooks, this book was like "streams in the desert." After finishing its delightful read, this thought came to the forefront in my mind, "If only high school textbooks would teach mathematics this way!" The book had serious limitations in the textbook context. Professor Péter only wanted to explain the exquisite structure of mathematics; her purpose was not to teach anyone mathematical skills or techniques. Despite these limitations, this book warrants reading by every math teacher and math student.

For about two decades I have toyed with the idea of writing a textbook based upon her approach. You have in your hands my attempt at doing this. I have enhanced many of her ideas while, based on my teaching experience and research, added much, much more. I hope what I have written has followed the same delight of mathematical exposition as exhibited by Professor Péter. By it, I give honor to both her memory and insight, a wonderful woman known affectionately by her students as Aunt Rózsa.

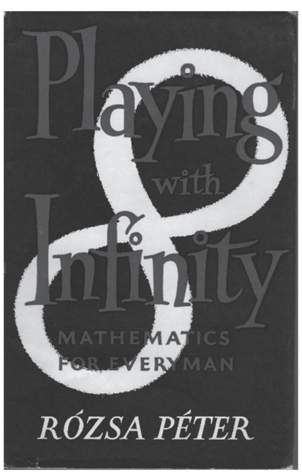

Figure 9. *Playing With Infinity*, G. Bell and Sons Ltd. edition.

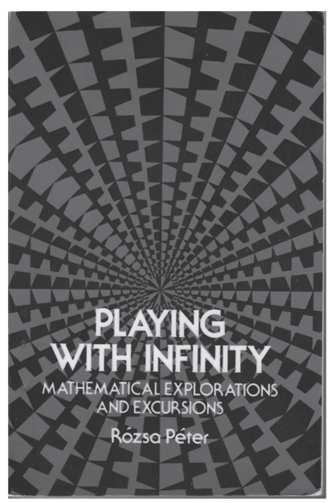

Figure 10. *Playing With Infinity*, Dover Publications edition.

---

[6] As a relatively new math teacher, I wanted to learn as much as I could from master teachers. Her title appeared in a book catalogue that I had ordered. Discovering Péter was like discovering gold.