Journal of Geoscience Education, v. 47, n. 2, p. 166-175, March1999 (edits, June 2005)

Computational Geology 5

If Geology, Then Calculus

H.L. Vacher, Department of Geology, University of South Florida, 4202 E. Fowler Ave., Tampa FL 33620

Introduction

This column is the fifth in the series, and so it closes the first year of "Computational Geology." I would be remiss if I let a year go by without discussing or giving an example of how calculus informs the practice of geology, or vice versa. Therefore, I will devote this column to the intersection of geology and calculus. Rather than discuss a computational problem illustrating the intersection, I shall try to prove the following assertion: "If you understand geology, you understand calculus."

Examining the Proposition

If you have had a course such as *Critical Thinking* or freshman logic, you will recognize the statement "If you understand geology, you understand calculus" as a compound proposition of the form, $p \rightarrow q$ (read: "If p, then q"). The p and the q each stand for an elementary (i.e., noncompound) proposition, where a *proposition* is defined as a statement that can assume either of only two values (T or F, called *truth values*; i.e., a proposition cannot have the truth value of "maybe"). In the case here, the two elementary propositions are:

p = "you understand geology"
q = "you understand calculus" (1)

The "if-then" statement is called a *conditional*. The p-statement is the *antecedent*; the q-statement is the *consequent*.

Just as the p and q can be only T or F, a conditional ($p \rightarrow q$) can be only T or F (i.e., again there is no "maybe"). The conditional is said to be F when the antecedent (p) is T and the consequent (q) is F. For any other circumstance, $p \rightarrow q$ is considered to be T. Thus, $p \rightarrow q$ is T if both p and q are T, and it is also T if p is F (whereupon q can have either value, T or F).

Two important Laws of Logic concerning conditionals are:

$$(p \rightarrow q) \leftrightarrow \sim (p \wedge \sim q) \quad (2)$$
$$(p \rightarrow q) \leftrightarrow \sim q \rightarrow \sim p \quad (3)$$

The symbols in these expressions indicate the following: "↔" is equivalence (read: "same as"); "~" is negation (read: "not"); and "∧" is conjunction (read: "and"). For completeness, there are two other symbols used in formal logic: "∨" for inclusive disjunction (read "either ... or ... or both"); and "⊻" for exclusive disjunction (read: "either ... or ... but not both"). With the translations of the symbols, statement (2) becomes: "Saying 'If p then q' is the same as saying 'It is impossible to have p and not have q'", and statement (3) becomes "'If p, then q' is the same as 'If not q, then not p'" The Law of Logic given in statement (3) is called the Law of Contraposition. The statement "$\sim q \rightarrow \sim p$" is the *contrapositive* of "$p \rightarrow q$ ".

With these translations and equivalences in mind, we can restate the proposed assertion, "If you understand geology, you understand calculus". With the equivalence of statement (2), the assertion becomes "It is not possible for you to understand geology and not understand calculus". With the equivalence of statement (3), the assertion becomes "If you do not understand calculus, you do not understand geology." All three of these statements are equivalent; if I prove one, I prove them all.

To examine the proposed assertion a little further, it might be useful to apply the words "sufficient" and "necessary", which often arise when discussing conditionals. In the conditional, $p \rightarrow q$, p is a *sufficient condition* for q, and q is a *necessary condition* of p. The term "sufficient condition" means that the condition is enough to bring about q. The term "necessary condition" means that the condition is unavoidable if p is present. This is the same as saying "p only if q"; in other words, saying "If p, then q" is the same as saying "p only if q".

With this language in mind, we can restate the proposed assertion again. New, equivalent statements are: "Understanding geology is a sufficient condition for understanding calculus," "understanding calculus is a necessary condition of understanding geology," and "you understand geology only if you understand calculus."

Before proceeding with the proof, perhaps I should clarify something. Although I am saying, "There is no one who understands geology who does not understand calculus", I am well aware that there are many people – professionals as well as students – who would claim that they are perfect counter-examples. These people would say: "Here I am – I understand geology but do not have a clue about calculus; moreover, I have the grades to prove it." Well, I do not accept that proof. My statement for q is you *understand* calculus. I am not arguing: "If you understand calculus, you know how to differentiate and integrate". To those of you who claim to be counter-examples, I say – You know more about calculus than you think you know. Don't let the skills of differentiating and integrating get in the way of concepts. My point is simple – and here I divulge my proof – there are some things about geology that guarantee an instinctive understanding of calculus.

One more thing: I am not arguing "If you understand calculus, you understand geology." That would be $q \rightarrow p$, which is the converse of $p \rightarrow q$. As they say, "The converse does not necessarily follow." In fact, it is easy to think of counter-examples to $q \rightarrow p$: people who understand calculus but do not understand geology. Lord Kelvin (1824-1907), one of the foremost mathematical physicists of his time, is a well-known historical figure who clearly understood calculus. Among other notable achievements, he proved from first principles of heat conduction that the Earth could not be as old as the Uniformitarians claimed. More importantly, he believed his proof. Clearly, calculus is not sufficient to understand geology.

Method of Proof

How to Solve It (Princeton University Press, 2nd ed., 1957, reprinted 1988) is a useful little book sometimes available at commercial bookstores. The book was written by G. Polya (1887-1985), who wrote several books on problem solving and mathematics education. The book aims to develop a "modern heuristic", which, according to Polya (p.129), "endeavors to understand the process of solving problems, *especially the mental operations typically useful* in this process" (emphasis is Polya's).

Polya outlined four steps to solving problems: (1) understanding the problem; (2) devising a plan; (3) carrying out the plan; and (4) looking back (including checking the result). We will go through these four steps in solving our particular problem, namely constructing a proof for "If geology, then calculus."

"Understanding the problem" is covered in the preceding section, "Examining the Proposition" – except for one thing. Under "understanding the problem", Polya wants us to draw a figure. This can be satisfied by Figure 1, which represents our assertion. The rectangle represents all people, and it is divided into two parts: all the people who do not understand calculus are on the left, and all the people who do understand calculus are on the right. The circle included within the right-hand side represents all the people who understand geology. Clearly, if a person understands geology, that person understands calculus – or at least that is what the figure shows. The equivalent propositions are shown equally well; for example, a person cannot both understand geology and not understand calculus.

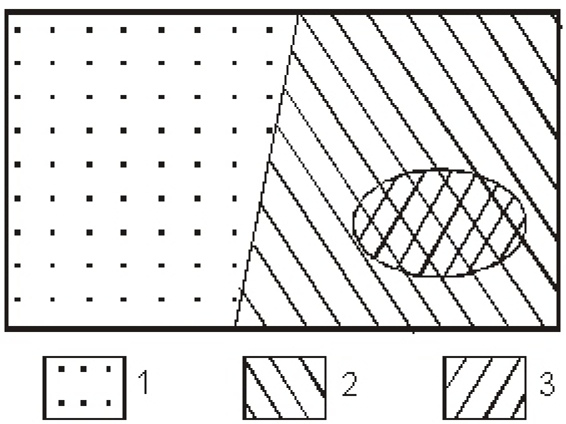

Figure 1. "If *p* then *q*." Key: 1 = ~*q*, 2 = *q*, 3 = *q*.

To prove the assertion, I will use the following argument structure:

P1. $p \rightarrow r$
P2. $r \rightarrow q$

C. $p \rightarrow q$ (4)

where "P1" and "P2" signal Premise 1 and Premise 2, respectively; "C" labels the Conclusion; the line indicates that a conclusion is drawn from the premises. This argument can be written on one line:

$$[(p \rightarrow r) \wedge (r \rightarrow q)] \rightarrow (p \rightarrow q) \qquad (5)$$

which says: "If it is true that *p* gets *r* and that *r* gets *q*, then it follows that *p* gets *q*". This is a well-known Law of Logic, the *hypothetical syllogism*.

The argument structure of (4) is *valid*. (All Laws of Logic are valid.) Saying that (4) is valid means that it is impossible for the premisses to be true and the conclusion false; said another way, if the premisses are true, validity of the argument guarantees that the conclusion is also true.

It is important to note that the preceding paragraph does not say that the validity of an argument guarantees that the conclusion *is* true. The guarantee is that the conclusion will be true *if* the premisses are true. If the premisses are not true, there is no guarantee; the conclusion may be true, or it may be false. If you want a guarantee that the conclusion is true, you need a *sound* argument. A sound argument is a valid argument with true premises.

There are formal – mathematical – ways of showing that an argument is valid, and if you have had *Critical Thinking* you have studied them. It is not necessary to trot out those methods here to convince you that the argument structure of (4) is valid. It is self-evident. Or, see Figure 2. The more important question – the one on which my proof will hinge – is whether I can come up with a sound argument of the form of argument structure (4) producing "If geology, then calculus."

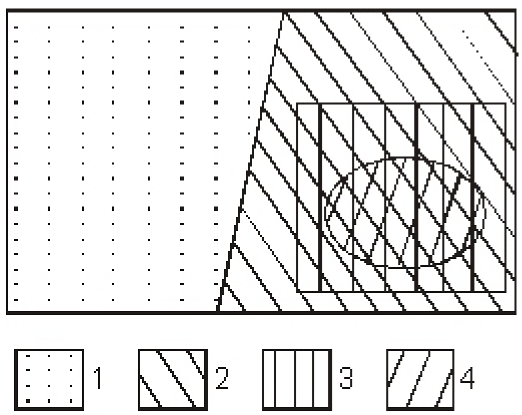

Figure 2. "If *p* then *r*" and "If *r* then *q*."
Key: 1 = ~*q*; 2 = *q*; 3 = *r*; 4 = *p*.

For example, consider the argument resulting from the substitution, *r* = "you understand snakes". Then, with the substitutions of (1), the argument of (4) becomes:

P1. If you understand geology, you understand snakes.
P2. If you understand snakes, you understand calculus.

C. If you understand geology, you understand calculus.

The argument is unquestionably valid, for it is simply a special case of (4). But the argument is not sound, because neither of the premises is true: understanding geology is not sufficient to understand snakes, and understanding snakes does not mean you understand calculus. With one or more false premises, a valid argument can result in a conclusion that is either true or false. In this case, I believe, the argument results in a true conclusion. But the argument does not prove the conclusion, because it is not sound. There is a crucial difference between valid and sound arguments.

The question, then, becomes: Is there a substitution for *r* that makes both P1 and P2 true. I believe there is: let *r* be "***you understand rates and maps***." Then the argument is:

P1. If you understand geology, you understand rates and maps.
P2. If you understand rates and maps, you understand calculus.

C. If you understand geology, you understand calculus.

This argument, I propose, is sound. In other words, to convince you that geology gets you an understanding of calculus, all I have to do is convince you that understanding rates and maps is a necessary consequence of understanding geology, and that understanding calculus follows from understanding rates and maps.

Understanding Calculus: Carrying Out the Plan.

What do I mean by "You understand calculus"? I believe it comes down to three questions: (a) Do you know what a derivative is? (b) Do you know what an integral is? (c) Do you know (the Fundamental Theorem of Calculus) that finding a derivative and finding an integral are inverse processes? While a geologist may not know the terms, I believe a geologist knows these things intuitively – from geological experience – because a geologist (a) understands topographic slopes, (b) understands volumes as portrayed on topographic maps, and (c) applies a form of uniformitarianism when, upon looking at the Grand Canyon, asks about the sediment load of the Colorado River. We will take them one at a time.

1. Topographic Slopes.

The derived function. Remember in your very first geology course you had a lab exercise in which you drew a profile of the land surface from a topographic map? Let's say for the sake of discussion that you generated the topographic profile shown in Figure 3: a topographic rise to a capstone-supported topographic high (such as a mesa). The question now is, How does the *slope* of the profile vary from one side of the profile to the other?

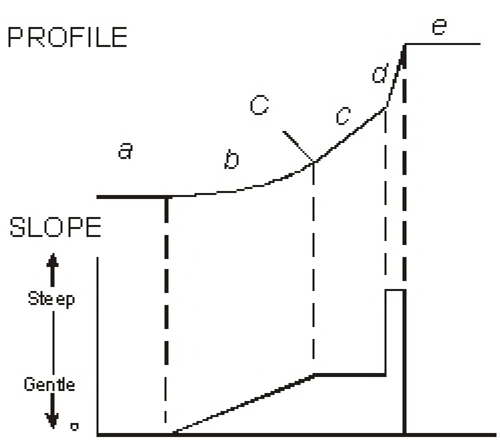

Figure 3. A topographic profile and its derived function (slope).

First, we need to review the meaning of "slope." The core of the definition familiar to geology students is *rise over run*; for example, if you are interested in the slope between two points on the ground, the slope is the difference in elevation between the points (the rise) divided by the horizontal distance between the points (the run). There is some divergence in usage in how the slope is expressed. Some people simply give the result of dividing the rise by the run (usually a decimal fraction); some take the arctangent of that number to produce the angle. Both quantities are legitimately called slopes, but here I will use *slope* for the rise over the run itself, and *slope angle* for the arctangent of the rise over run. If you multiply the slope by 100, you produce a percent, the *grade*. Scenic highways with 7% grade, for example, commonly merit a warning sign; the corresponding slope is 0.07, and the slope angle is 4^{o} (although they certainly seem steeper than that, at least to me).

Now, let's return to the variation in slope across the topographic profile of Figure 3. I bet you would describe the slopes at the localities indicated by the letters something like this: the slope at *a* is zero (a horizontal surface); through segment *b*, the slope increases and merges with segment *c,* where the slope is constant at a moderate value; then the slope abruptly increases to a high value in the cliff at *d*; finally, at the top, the slope is zero again, in *e*. The graph below the profile shows the variation that has just been described (neither the graph nor the profile is drawn to scale in Fig. 3; they are for diagrammatic purposes).

Here is the point to my argument. The topographic profile is a function, $f(x)$: for every distance (x) from the starting point on the left side of the profile, there is an elevation (represented here by $f(x)$. The elevation $f(x)$ is a continuous function. The rate that $f(x)$ increases is shown by the lower graph. This curve, too, shows a function. This function, in the language and notation of Joseph Louis Lagrange (1736-1813), is called the *derived function* and denoted by $f'(x)$. For every distance (x) there is a value of the slope, $f'(x)$. This derived function, $f'(x)$, is also known as the derivative of $f(x)$. See, geologists know, and feel, what a derivative is: it is the slope function. As geologists walk around the topography, they experience the slope function.

Fermat's ratio. To further convince you have an instinct for derivatives, suppose you were located on the ground somewhere in the segment *b*: How would you measure the slope there? It might take some thought, but after field camp, you could come up with something like the following. Suppose the vertical distance from the bottom of your boots to your eye level is 1.7 m (when you stand straight up). Then set the clinometer of your Brunton compass to zero; mark your position on the ground; sight horizontally through your Brunton to identify the spot where your line of sight intersects the ground; go mark that position; and measure the distance between the two marked spots. Let's say that the distance is 24.4 m (on the ground). What you have, then, is two sides of a right triangle: the rise is the height of your eye level (1.7 m), and the hypotenuse is the measured distance (24.4 m). The run is the horizontal leg of the triangle and can be found from the Pythagorean relation:

$$run = \sqrt{24.4^2 - 1.7^2} = 24.34 \ , \qquad (6)$$

in m, with one extra digit. Then the slope is:

$$slope = \frac{rise}{run} = \frac{1.7}{24.34} = 0.07\,, \tag{7}$$

corresponding to a slope angle of 4º (and a grade of 7%).

The measurement of the slope as just described is actually a hillside version of the approach Pierre de Fermat (1601-1665) used to calculate the slope of the tangent to curves represented by polynomials. Fermat's approach was to calculate the ratio

$$ratio_{Fermat} = \frac{f(x+a) - f(x)}{a}\quad , \tag{8}$$

where a is a *little bit* added on to x. In our case on the hillside, a is 24.34 m, the run. The difference between the elevation of the hillside at $x + a$ and the elevation of the hillside at x is the rise (1.7 m), or the numerator in the equation. Clearly, Fermat's ratio is the rise over run, or the slope.

It is important to realize that the slope represented by Fermat's ratio is almost the slope of the *tangent*, the straight line that touches the curve at only the point x. The relation between Fermat's ratio and the tangent can be seen by examining the curve in Figure 4, which, for

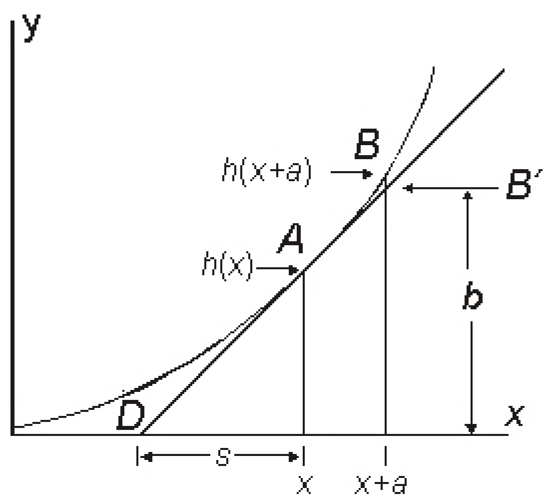

Figure 4. A line tangent to a curve to illustrate Fermat's ratio. Adapted from C.H. Edwards, Jr., *The Historical Development of the Calculus*, Springer Verlag, New York, 1979. Chapter 5, Figure 3.

geological familiarity, we can say is a hillside; therefore, h represents elevation, and x is the horizontal distance from some origin. Point A is the point of interest; its elevation is $h(x)$ – read, "h at x". Point B, also on the hillside, is a little further away -- a distance $x + a$ from the origin; the elevation of B is $h(x+a)$, or "h at $x + a$". The tangent touches the hillside at x and is extended downward to the x-axis; the intersection of the tangent with the x-axis is a horizontal distance s from A. Meanwhile, the tangent is extended in the other direction and, at B', intersects the vertical line dropped down from B. The elevation of B' is b. Now, here's the key: by similar triangles,

$$\frac{s+a}{s} = \frac{b}{h(x)} . \tag{9}$$

Fermat recognized that, using our notation, b is nearly the same as $h(x+a)$, and so by substituting $h(x+a)$ for b in Equation 9,

$$\frac{s+a}{s} = \frac{h(x+a)}{h(x)} , \tag{10}$$

from which by rearranging,

$$\frac{h(x+a) - h(x)}{a} = \frac{h(x)}{s} . \tag{11}$$

The left side of Equation 11 is Fermat's ratio; the right side is the slope of the tangent to the curve at x. Therefore, Fermat's ratio is the slope of the tangent, if b is taken to be $h(x+a)$.

To see how Fermat's ratio works on a polynomial, we can take $h(x) = mx^2 + C$, where m is an arbitrary coefficient and C is an arbitrary constant:

$$ratio_{Fermat} = \frac{\left[m(x+a)^2 + C\right] - \left[mx^2 + C\right]}{a} , \tag{12}$$

Expanding $(x+a)^2$ produces:

$$ratio_{Fermat} = \frac{\left(mx^2 + 2mxa + ma^2 + C\right) - \left(mx^2 + C\right)}{a} . \tag{13}$$

The mx^2 terms subtract out, and so does the C (an important point later), and so we are left with:

$$ratio_{Fermat} = \frac{m\left(2ax - a^2\right)}{a} . \tag{14}$$

Now, because a is very small relative to x, the a^2 term can be dropped as negligible next to $2ax$. The end result is:

$$ratio_{Fermat} = 2mx . \tag{15}$$

Fermat was one of the seventeenth century mathematicians who anticipated the invention of calculus by Isaac Newton (1642-1727) and Gottfried Wilhelm Leibniz (1646-1716). Newton and Leibniz are credited with the invention because they, among other things, realized that the tangent needs to be thought of in terms of infinitely small quantities (infinitesimals) and general rules (not just *ad hoc* problems such as polynomials). In modern language, Fermat's ratio is a finite difference, not strictly a derivative. Newton spoke of a ratio of "evanescent quantities", the infinitesimals. The concept of a ratio of quantities that don't exist greatly exercised later philosophers, such as George Berkeley (1685-1753), as illustrated by his 1734 essay, *The Analyst, or a Discourse Addressed to an Infidel Mathematician*. The infidel, incidentally, was Edmund Halley (1656-1742), a disciple of Newton.

The limit concept. Looking again at Figure 4, it is quite evident that the rise over the run using the points on the ground (A and B) is different from the rise over the run as calculated between A and B'. In the first case, the rise over the run is the *average slope between A and B*. In the second case, the rise over the run is *the slope at A*. They are not quite the same because of the vertical difference between B and B'.

So, if the geologist really needed the slope at A – i.e., if the average slope between A and B were not good enough – what would the geologist do? Recall in the earlier example (with the 4° slope angle) the vertical distance between A and B was determined by the geologist's eye-height, and so the horizontal distance was substantial – 24 m. The geologist would instinctively use smaller and smaller distances. Then, as the distance a becomes smaller, B and B' become more nearly the same, and the ratio of rise over run becomes more nearly the slope of the tangent at A. In other words, in the limit as a goes to zero, Fermat's ratio (rise over run) becomes the slope of the tangent. This is what geologists measure when they kneel down at A and measure the angle with the clinometer of the Brunton. An alternative (and less gymnastic) way of measuring it would be to stretch a taught string between two points very close together on either side of A and measure the rise over the run along the string using a plumb bob and level.

Thus the slope at a point – the derivative – is not a ratio of quantities that do not exist, but rather the limit of a ratio. This concept, which was implied in the work of Newton and Leibniz – and which is implied when the geologist takes smaller and smaller excursions from A to measure the slope at A – was formalized by Augustin-Louis Cauchy (1789-1857) decades after Lagrange. It was then that calculus was put on a rigorous footing and the derivative was defined in the way that it is taught today in calculus courses:

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} \qquad (16)$$

The differential coefficient. Geology students who are familiar with slopes are instinctively in step with Leibniz. It was Leibniz who wrote the slope of the tangent as dy/dx – more than 100 years before Lagrange came up with $f'(x)$ for the "derived function" (derivative) and Cauchy defined it as a limit. Before "derivative", the ratio dy/dx was known as the "differential coefficient."

For an explanation of the Leibniz notation (the d-terms), there is none better than that given by Silvanus P. Thompson (1851-1916, contemporary with Kelvin) in an irreverent, and extremely helpful, little book: *Calculus Made Easy* (St. Martin's Press, Third Edition, 1946, reprinted 1987). The subtitle of Thompson's book is appropriate: *Being a Very Simplest Introduction to those Beautiful Methods of Reckoning which are Generally Called by the Terrifying Names of the Differential Calculus and the Integral Calculus*. Chapter 1, "To Deliver You from the Preliminary Terrors", starts with the following (p. 1).

> The preliminary terror, which chokes off (students) from even attempting to learn how to calculate, can be abolished once for all by simply stating what is the meaning -- in common-sense terms -- of the two principal symbols that are used in calculating.
> These dreadful symbols are:
> (1) d which merely means "a little bit of'.
> Thus dx means a little bit of x; or du means a little bit of u. Ordinary mathematicians think it more polite to say "an element of", instead of "a little bit of". Just as you please. But you will find that these little bits (or elements) may be considered to be indefinitely small.

We will get to the second "dreadful symbol" shortly, but for now we should skip to Chapter 3 ("On Relative Growings"), where Thompson gives the key concept of the differential coefficient. From *Calculus Made Easy*, p. 8:

> Suppose we have got two ... variables that depend one on the other. An alteration in one will bring about an alteration in the other, *because* of this dependence. Let us call one of the variables x, and the other that depends on it y.
> Suppose we make x to vary, that is to say, we either alter it or imagine it to be altered, by adding to it a bit which we call dx. We are thus causing x to become $x+dx$. Then because x has been altered, y will have altered also, and will have become $y+dy$. Here the bit dy may be in some cases positive, in others negative; and it won't (except very rarely) be the same size as dx.

From *Calculus Made Easy*, p. 10:

> Now right through the differential calculus we are hunting, hunting, hunting for a curious thing, a mere ratio, namely, the proportion which dy bears to dx when both of them are indefinitely small.
> It should be noted ... that we can only find this ratio dy/dx when y and x are related to each other in some way, so that whenever x varies y does vary also."

Thompson's explanation is accompanied by a diagram like Figure 5 showing a little right triangle, with one leg having a length dy and the other leg having a length dx. This triangle is the same as used by Leibniz, who called it *the characteristic triangle*. A geologist, thinking of y as elevation, would see the dy as the rise and the dx as the run at the site with a location coordinate of x and an elevation of y. Thus, dy/dx is simply rise over run, or the slope at the location given by x.

So, for the purpose of my argument: if you understand geology, you understand slopes, which means you understand rise over run, which means you understand Leibniz's df/dx. You also understand that the slope at a point, expressed as df/dx or $f'(x)$, can only be approximated with the ratio of a finite Δf over a finite Δx; that is, it is only when the Δx becomes infinitely small (infinitesimal) that the ratio of finite differences (average slope through the distance Δx) becomes the actual slope (derivative at x).

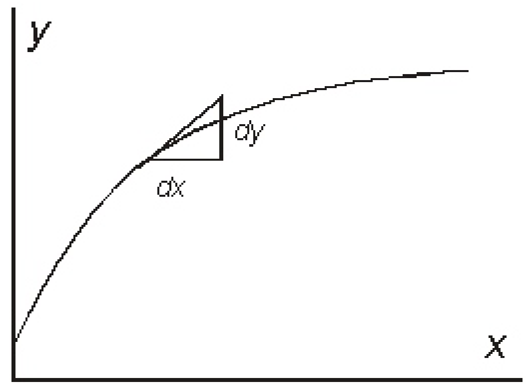

Figure 5. The characteristic triangle. Adapted from C.H. Edwards, Jr., *The Historical Development of the Calculus*, Springer Verlag, New York, 1979. Chapter 9, Fig. 3.

2. Areas and Volumes

Another exercise from the lab of your introductory geology course no doubt involved visualizing three-dimensional landforms portrayed by the contours of topographic maps. For example, how do you visualize the idealized island shown in Figure 6? Specifically, imagine that you were asked to make a three-dimensional, scale model of the island. What would you do?

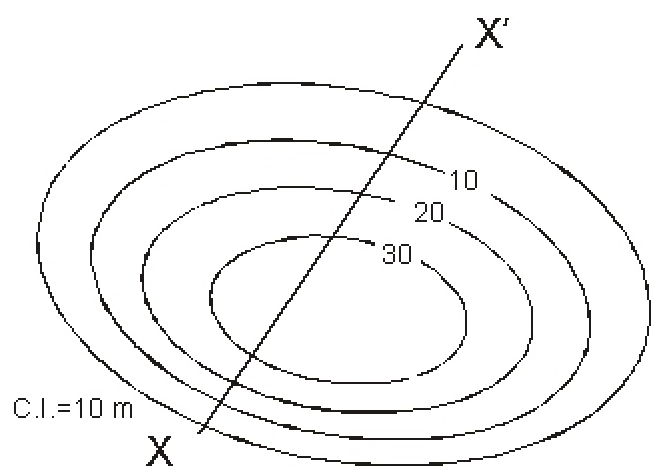

Figure 6. Topography of a hypothetical island.

Many students answer – and some lab books illustrate – that one can trace out the contours onto cardboard, cut out the areas enclosed by the contours, and stack them up. The model would look something like a wedding cake, a stack of slabs, each slab representing a thickness equal to the contour interval of the map.

Now, suppose you were asked to determine the volume of rock comprising the island (above sea level). Remembering the wedding cake model and thinking how you would determine its volume, your familiarity with maps can lead to a very straightforward approach: (a) measure the area of each of the slabs and, using the scale of the map, convert the model areas to actual areas; (b) multiply those areas by the thickness represented by the slab (contour interval) to get the volume of each slab of rock; and (c) add up those volumes. Knowing the difference between the terraced topography of a wedding cake and the smooth slopes implied by Figure 6, however, you would also know that this sum-of-slabs approach would yield only an approximation. This approximation is:

$$V \approx \sum_{i=1}^{n} A_i \Delta h \, , \tag{17}$$

where A_i is the area of the i'th slab of rock, n is the number of slabs, and Δh is the contour interval. How could you make the approximation more accurate? The answer is easy: By having more contours and a smaller contour interval. In the extreme, you would want an infinitude of contours and an infinitesimal contour interval.

You said you would measure the area. How would you do that? That is straightforward too, and you have probably done something like this several times: trace the contour on graph paper; count the squares (let's call them rectangles, to be more general); and multiply the number of rectangles by the area represented by each rectangle. In other words, you would find

$$A \approx \sum \Delta x \Delta y \, , \tag{18}$$

where Δx is the length represented by one side, Δy is the length represented by the other side, and $\Delta x \Delta y$ is the area represented by each rectangle, Anyone who has ever counted squares knows that in order to properly account for the edges, you need mini-squares -- the smaller the better, for a better approximation. In the extreme you want an infinitude of infinitesimal rectangles.

We can now return to Thompson's comments on the two "dreadful symbols" of calculus. Recall, the first was d for *little bit*. The second is (*Calculus Made Easy*, p. 1-2):

> ... $\int$ which is merely a long S, and may be called (if you like) 'the sum of'
>
> Thus $\int dx$ means the sum of all the little bits of x; or $\int dt$ means the sum of all the little bits of t. Ordinary mathematicians call this symbol 'the integral of'. Now any fool can see that if x is considered as made up of all of little bits, each of which is called dx, if you add them all up together you get the sum of all the dx's (which is the same thing as the whole of x). The word 'integral' simply means 'the whole'. If you think of the duration of time for one hour, you may (if you like) think of it as cut up into 3600 little bits called seconds. The whole of the 3600 little bits added up together make one hour.
>
> When you see an expression that begins with this terrifying symbol, you will henceforth know that it is put there merely to give you instructions that you are now to perform the operation (if you can) of totalling up all the little bits that are indicated by the symbols that follow.
>
> That's all.

And so ends Thompson's Chapter 1!

The combination of symbols $\int dx$, like dy/dx, is from Leibniz. Using this notation for summing up all the little bits, Equation 17 becomes

$$V = \int A dh \quad . \tag{19}$$

Note that the *approximately equals* sign has been replaced by an *equals* sign. The understanding is that there is an infinitude of contours and an infinitesimal contour interval (specifically dh). Similarly, Equation 18 becomes

$$A = \iint dx dy \ . \tag{20}$$

with the understanding that a finite number of finite rectangles has been replaced by an infinitude of infinitesimal ones. Thus the integral is the limit of a sum, just as the derivative is a limit of a ratio.

To further convince you that the integral is nothing but a sum that you instinctively know from map experience, consider one last problem. Suppose you are consulted by someone who wants to dig a sea-level canal from one side of the island to the other. Leaving aside environmental questions, suppose the canal is to be 50 ft wide and extend along line X-X' of Figure 6. How much material needs to be removed (ignoring the below-sea-level part)?

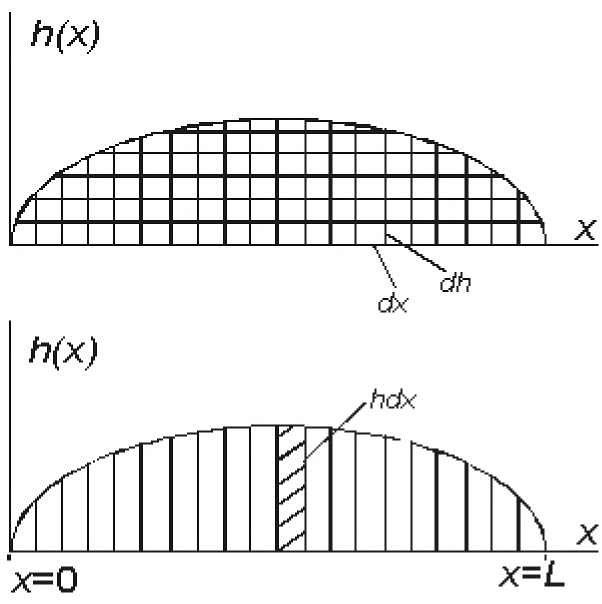

Figure 7. Determining the area using tiny squares as compared to slim columns.

Clearly, again, you would use the map to draw a cross profile (on graph paper), and then you would determine the area below the curve (Fig. 7) and multiply the 50-ft width to get the volume. You might start by counting tiny squares. However, after counting a few zillion squares – if you hadn't already thought of it – you would begin to think of the network of squares as a collection of side-by-side columns each having the width of one square (maybe remembering basalt columns). Then, rather than counting all the squares, you would simply multiply the height of the column (read from the graph paper) by the width of a square to determine the area of the column. Again, letting the height be *h*, the area of the column would be *hdx*, where *x* is the length of the canal and *dx* is the "little bit" of that length. Then, adding up all the columns, the total area of the cut across the island is:

$$A = \int_0^L h(x)dx \ , \tag{21}$$

So, although you may not use the notation, in splitting, measuring, multiplying and then summing, you are integrating.

3. Continental Denudation

The Fundamental Theorem. Sir Alexander Geikie (1835-1924), a contemporary of Kelvin, was professor at the University of Edinburgh, director-general of the Geological Survey of Great Britain and Ireland, and a champion of Uniformitarianism. He coined the phrase, "The present is the key to the past." Of interest here is a calculation he performed while arguing for the recency of geomorphic features such as lakes and valleys. The calculation appears in his "On Denudation now in Progress" (Geological Magazine, vol 5, p. 249-254, 1868; excerpted in K. Mather and S.L. Mason, *A Source Book in Geology*, Harvard University Press, p. 523-527, 1939):

> The extent to which a country suffers denudation at the present time is to be measured by the amount of mineral matter removed from its surface and carried into the sea. An attentive examination of this subject is calculated to throw some light on the vexed question of the origin of valleys and also on the value of geological time....
>
> ....
>
> Comparing the measurements which have been made of the proportion of sediment in different streams we shall probably not assume too high an average if we take that of the carefully elaborated Survey of the Mississippi. This gives an annual loss over the area of drainage equal to 1/6000 of a foot. If then a country is lowered by 1/6000 of a foot in one year, should the existing causes continue to operate undisturbed as now, it will be lowered ... 100 feet in 600,000 years, and 1000 feet in 6,000,000 years. The mean height of the Continents ... is in Europe 671, North America 748, South America 1151, and Asia 1132 English feet. Under such a rate of denudation therefore Europe must disappear in little more than four million of years, North America in about four millions and a half, South America and Asia in less than seven millions. These results do not pretend to be more than approximative, but they are of value inasmuch as they tend to shew that geological phenomena, even those of denudation, which are often appealed to as attesting the enormous duration of geological periods, may have been accomplished in much shorter intervals than have been claimed for them.
>
> ....
>
> "It seems an inevitable conclusion that those geologists who point to deep valleys, gorges, lakes, and ravines, as parts of the primeval architecture of a country, referable to the upheavals of early geological time, ignore the influence of one whole department of natural forces. For it is evident that if denudation in past time has gone on with anything like the rapidity with which it marches now, the original irregularities of surface

produced by such ancient subterranean movements must long ago have been utterly effaced."

Geikie's arithmetic is straightforward:

$$\text{Amount}_{\text{denudation}} = (\text{Rate}_{\text{denudation}})\,(\text{Amount}_{\text{time}}) \,. \qquad (22)$$

His problem was to find the amount of time given a denudation rate and the amount of denudation needed to erase a continent. Using his numbers for Europe:

$$671 \text{ ft} = (0.000167 \text{ ft/yr})(4.026\text{x}10^6 \text{ yr}) \,. \qquad (23)$$

This arithmetic is shown on a graph in Figure 8, with denudation *rate* plotted against time. The denudation rate is constant at 0.000167 ft/yr for a little more than 4 million years. The important thing is the area of the rectangle under the curve: it is given by the product on the right-hand side

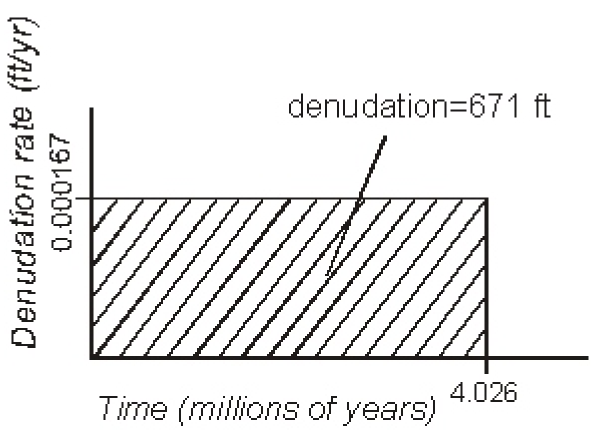

Figure 8. Denudation rate vs. time for a constant rate.

of Equation 23. In other words, the *amount* of denudation is given by the *area under the curve showing denudation rate vs time*. The denudation rate is rate of change of h (continental height) with respect to time. Letting h = continental height, the rate of change of h with respect to time is dh/dt, the derivative of h with respect to time. So finding h amounts to finding the *area under the curve of* dh/dt *vs time*. The arithmetic of Geikie illustrates the fundamental theorem of calculus in its simplest form.

Geikie noted that he was assuming that "the causes continue to operate undisturbed as now". What if he had assumed they changed from time to time? To answer that question, we can return to Figure 3.

Suppose you are a geologist standing at location *C*. Suppose the elevation there is 3,025 ft. You climb to the top of the platform at *e*, keeping track of your rate of ascent. In *c*, you climb at the rate of 72 ft per 1000 ft (0.072) in a *horizontal* distance of 12,000 ft; in *d* you climb at the rate of 325 ft per 1000 ft (0.352) in a horizontal distance of 2,400 ft. What is your elevation when you get to *e*?

As a geologist, you would quickly respond something like the following: "Starting at 3,025 ft, I climbed 72 ×12 or 864 ft in section *c* to reach an elevation of 3,025+864 or 3,889 ft, and then I climbed 325 × 2.4 or another 780 ft to reach a final elevation of 4,669 ft" (one hopes rounding off to 4700 ft to allow for uncertainties – see CG-1, May 1998).

Your response is shown graphically in Figure 9. The upper graph shows the slopes – the rates of change of elevation. The areas under the two slopes define elevation vs distance in the lower graph. Clearly, finding areas in Figure 9 reverses the operation of finding slopes in Figure 3. Figure 9 also shows that when the slope changes and you need to find the elevation changes piecemeal, you are finding partial areas and summing them. Each one of the partial sums can itself be viewed as the sum of a little bit of *x* times the value of the slope, as shown in Figure 10. In mathematical notation, Figure 10 can be expressed as:

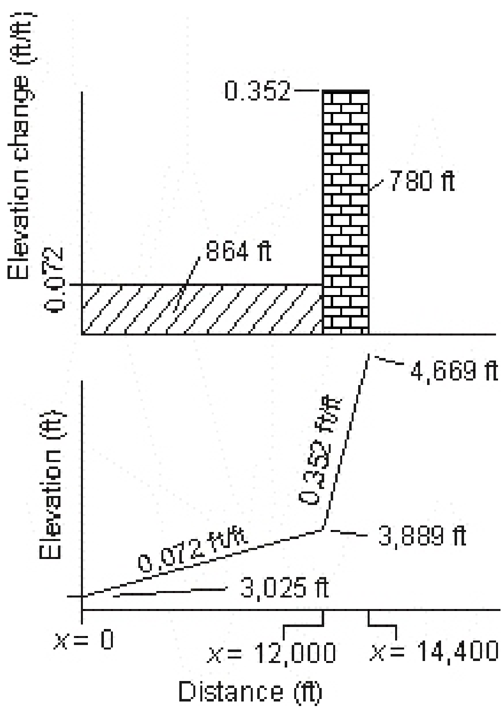

Figure 9. Rate of ascent vs. distance and the inferred elevation vs. distance.

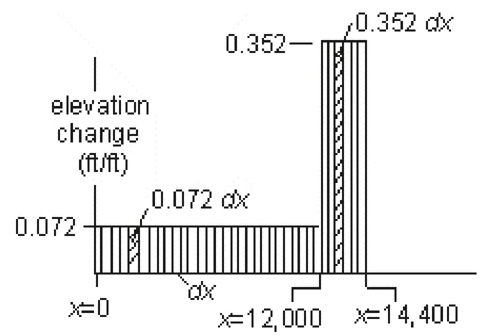

Figure 10. Summing columns to find the areas in Figure 9.

$$h_f = h_0 + \int_0^{12,000} 0.072dx + \int_{12,000}^{14,400} 0.352dx \quad , \qquad (24)$$

which shows, in addition to everything else, that after finding the area you add it to an initial value (representing the "constant of integration" of first-semester calculus).

Knowing that the numbers 0.072 and 0.352 are the topographic slopes, and that slopes are nothing more than derivatives, then you can see that Equation 24 is simply an example of:

$$h_1 = h_0 + \int_{x_0}^{x_1} \frac{dh(x)}{dx} dx , \qquad (25)$$

where $h = h_0$ at $x = x_0$ and $h = h_1$ at $x = x_1$. The integral is known as a definite integral because it represents a sum between definite limits.

As shown by Equation 25, integration is the reverse of differentiation. The same concept is shown by:

$$f(x) = \int \frac{df}{dx} dx + C , \qquad (26)$$

or

$$f(x) = \int f'(x) dx + C , \qquad (27)$$

in the Leibnizian and Lagrangian notations, respectively, and with *C* representing the constant of integration. The *C* is included because constants drop out in differentiation (recall our calculation of Fermat's ratio in Equation 12).

The integrals in Equations 26 and 27 are known as indefinite integrals. They are indefinite because the constant needs to be evaluated for particular applications. If the value of the function is known at x_0 , then $C = f(x_0)$. Then, Equation 27 becomes

$$f(x) = f(x_0) + \int_{x_0}^{x} f'(x) dx . \qquad (28)$$

To convince yourself of this, expand the definite integral in Equation 28 to get:

$$f(x) = f(x_0) + f(x) - f(x_0) . \qquad (29)$$

Equation 28 says that $f(x)$ is the area under the derived curve from an initial point, x_0 , to x, the point in question, added to the initial value, $f(x_0)$.

Integration and Antidifferentiation. A synonym for indefinite integral is *antiderivative*, a word that captures the concept of the fundamental theorem of calculus. A representative statement of the idea is given by J. Dantith and R. D. Nelson in the *Penguin Dictionary of Mathematics* (Penguin Books, London, 1989, p. 176):

> "If $F(x)$ is a function of x that, when differentiated, gives $f(x)$, then $F(x)$ is said to be the integral (or antiderivative) of $f(x)$ ...

Thus, integration can be thought of in three ways: (1) finding the area under a curve; (2) finding a sum of infinitesimal slices; and (3) finding the antiderivative. That all three processes result in the same answer can be illustrated by reconsideration of Figure 3.

In sector *b* of the profile shown in Figure 3, the *slope increases at a constant rate*. In sector *a* the slope is zero, and according to the problem we worked out with Figure 9, the slope is 0.072 ft/ft in sector *c*. As our new problem, let's suppose that sector *b* of Figure 3 is 7,200 ft long and that, within it, the *slope increases at the rate of 0.001 ft/ft per hundred feet*. How much does the elevation rise in sector *b*? In other words, if the elevation at *C* is 3,025 ft (as we said for the problem worked out in Fig. 9), what is the elevation in sector *a* of Figure 3, where the surface is horizontal?

The first two ways of solving the problem are shown in Figure 11. In each case, we plot the slope (dh/dx) against distance (x) from $x = 0$ to $x = 7{,}200$ ft. The area under the curve gives

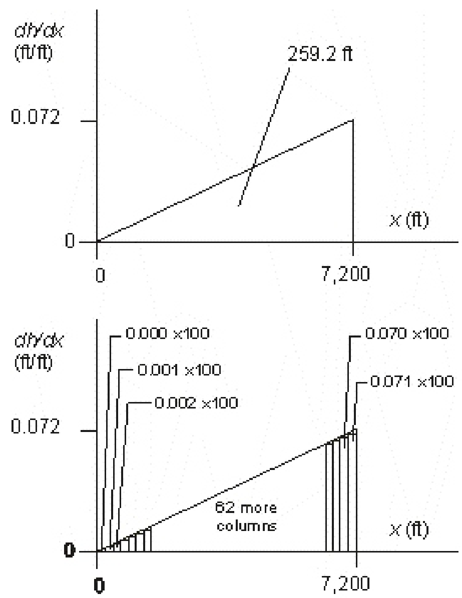

Figure 11. Finding the required area by using a geometric formula (top) or a sum of many columns (bottom).

the total change in h in the 7,200 ft. In the first case, the area (Δh) can be found simply from the formula for the area of the triangle

$$\begin{aligned}\Delta h &= \frac{(height)(base)}{2} \\ &= \frac{(0.072\,ft/ft)(7200\,ft)}{2} \\ &= 259.2\,ft\end{aligned} \tag{30}$$

In the second, the area can be found by summing 72 strips each 100 ft across. Taking the lower sum (i.e, the rectangles that lie under the curve, as shown in the Figure 11), the sum consists of the following 72 terms:

$$\Delta h = (0.000)(100) + (0.001)(100) + (0.002)(100) + \ldots + (0.071)(100)\,. \tag{31}$$

Removing the common factors (0.001 ft/ft and 100 ft), we get:

$$\begin{aligned}\Delta h &= 0.1\,(0+1+2+\ldots+71)\\ &= (0.1)\,\frac{(71)(72)}{2}\\ &= 255.6\,ft\end{aligned} \quad , \tag{32}$$

which makes use of the formula for the sum of the first n integers: $n(n+1)/2$. The result of Equation 32 is an underestimate because all the rectangular columns lie under the curve. We can get closer by having slimmer rectangles (and more of them); that is, we can take the limit of the sum. Alternatively, we can get an overestimate by using rectangles that form an upper sum of 72 terms:

$$\begin{aligned}\Delta h &= (0.001)(100) + (0.002)(100) + \ldots + (0.0072)(100)\\ &= (0.1)(1 + 2 + 3 + \ldots + 72)\\ &= 262.8 \text{ ft.}\end{aligned} \tag{33}$$

The average of these two 72-term estimates (from Equations 32 and 33) is 259.2 ft, the same as that calculated in Equation 30.

Finally, we can find Δh by antidifferentiation. The slope is given by

$$\frac{dh}{dx} = 0.00001x\,. \tag{34}$$

Now we simply need to recall: Have we ever differentiated a function to produce a constant times x? The answer is yes, several pages ago, in the discussion of Fermat's ratio (Equation 15). From that work, we know that the antiderivative of cx is $cx^2/2$, and so, using a definite integral,

$$\begin{aligned}\Delta h &= \int_0^{7,200} 0.00001x\,dx\\ &= 0.00001\left[\frac{7200^2}{2} - \frac{0^2}{2}\right]\\ &= 259.2\,ft\end{aligned} \quad . \tag{35}$$

The ease with which the integral is found in Equation 35 – and the generality of the method – illustrates why antidifferentiation is the preferred process by which integrals are evaluated. The key, however, is knowing derivatives, because integration is differentiation backwards. That is why you take *Differential Calculus* before *Integral Calculus*.

Looking Back and Concluding Remarks

The fourth step of Polya's problem-solving strategy is to review the calculation or argument, see whether it can be seen at a glance, and inquire whether the result makes sense. To review, my argument at a glance is that a geologist's understanding of rates and maps carries over and gives an instinctive understanding of fundamental concepts of calculus. When thinking of a topographic slope, the geologist thinks of rise over run at the point, and this is the dy/dx of Leibniz. When thinking of the area of a cross-section, or the volume of a landform, the geologist thinks of sums of tiny (infinitesimal) areas or volumes, and this is $\int\int dxdy$ or $\int ydx$ for the cross section, or $\int\int\int dxdydz$ or $\int\int zdxdy$ for the volume. When thinking of the total volume (V) of material eroded from a continent, the geologist thinks of the erosion rate (dV/dt) and integrates it over time to produce V, thereby applying the fundamental theorem of calculus.

Does "Geology → Calculus" make sense? I believe it does, if one is careful to parse *conceptual understanding* from *ability to execute*. In other words, I do not contend that the geologist's instinctive understanding of concepts of calculus gives the geologist any ability to *use* calculus. Recalling the subtitle of Thompson's book, we can say that calculus consists of various *methods of reckoning*. The ability to use calculus comes, in part, from learning these methods – and practice. The other part is the knack of abstracting idealized conceptualized models; this is not normally a problem for geologists, as evidenced by their propensity for simplified cartoons of geologic conditions and Earth processes.

Taking the position that geology – more so now than ever before – is the study of rates as applied to the Earth, then it seems reasonable to think that geology students would want to learn the methods of reckoning that deal specifically with rates. That means that you will want to know calculus. With calculus, for example, you leave the world of description and enter the world of analysis. Cartoon abstractions can give way to quantified, and therefore testable, process-response models.

Addendum, June 2005

A new edition of *Calculus Made Easy* is available: *Calculus Made Easy* by Sylvanus Thompson and Martin Gardner, St. Martin's Press, New York, 1998. The book is the same Thompson masterwork, with three new, preliminary chapters by Gardner, the extraordinary mathematics explicator best known for more than 25 years of "Mathematical Games" columns in *Scientific American*.

For more about Martin Gardner, see the two-part interview by Dan Albers in the May and September 2004 issues of *The College Mathematics Journal* on the occasion of Gardner's 90th birthday. Gardner's career includes more than sixty books. Subjects include magic, philosophy, literature and pseudoscience, as well as the monumental works in mathematics.

Here is one of the reviews of the new *Calculus Made Easy* posted at Amazon.com.

> "*Calculus Made Easy* is arguably the best math teaching ever. To a non-mathematician, its simplicity and clarity reveals the mathematical genius of Newton, Leibniz, and Thompson himself. Martin Gardner deserves huge thanks for renewing this great book."—*Julian Simon, author of Population Matters*