A Product Calculus

Michael Z. Spivey
University of Puget Sound
Tacoma, Washington 98416-1043
mspivey@ups.edu

The origin of this article is a conversation I had with a first-semester calculus student in the fall of 2005. We had just covered the idea of the derivative in class. I had taken the standard approach of motivating the derivative with the velocity and slope interpretations while trying to emphasize that it is a more general concept involving rate of change. One of my students was not satisfied with my explanations, though, and she and I had a few discussions outside of class on the subject. At one point we had a conversation that went something like this:

Sarah: I still don't understand the derivative. I know you said it gives you different things, like slope and velocity, depending on what you're talking about, but what does it really *mean*?

Spivey: Well, when you get down to it, it really means what its definition says. [*Writes the following on the board:*

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}.\Bigg]$$

Sarah: But what does *that* really mean?

Spivey: Well, you can think of it as a measure of how two quantities change relative to each other. [*Writes the following on the board:*

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}.\Bigg]$$

Sarah: But why do you *divide* the changes?

[*Good question. Purely from the "comparing two changes" standpoint, what is so special about dividing the changes?*]

Spivey: Um... because it gives us something useful. [*Pause.*] I suppose if you did something other than divide the changes you might get a different kind of calculus.

Sarah: Oh! [*Pause.*] Cool! [*Pause.*] You can *do* that?

I didn't know if you could do that or not; I was just guessing. It turns out, though, that a slight modification to the definition of the derivative does lead to a different kind of calculus, one that bears the same relationship to multiplication that the ordinary calculus does to addition. This *product calculus* dates to at least Volterra [9] in the late 1800's and is well-known in certain areas of mathematics in which it has applications (such as particular subfields of statistics [3] and differential equations [2]). However, despite its natural appeal, it does not appear to be well-known to the mathematical community as a whole. Most treatments of the subject are also written for experts in the area, making it difficult for those interested in learning more about product calculus to do so. In particular, there does not appear to be a comprehensive treatment of product calculus accessible to undergraduates. Some of the more accessible sources are a text by Dollard and Friedman [2], a survey by Gill and Johansen [3], a short article by Guenther [5], and a discussion on some of the message boards and elsewhere at the Math2.org web site [6, 7]. The Dollard and Friedman text is a good resource on the product integral, but it has very little on the product derivative. The Gill and Johansen survey also focuses almost solely on the product integral. It is aimed at statisticians, too, and nearly all undergraduates will have trouble with it. The Guenther article is too short to include much detail and has very little on the product derivative. The discussions at Math2.org do include some discussion of the product derivative, but they are (by nature) not organized and not complete.

The purpose of this article, then, is to provide an initial reference, at the undergraduate level, for those interested in learning more about product calculus. We define the product derivative, give some rules for calculating product derivatives, discuss a few applications of the product derivative, and work through some theorems about the product derivative necessary for a proof of a fundamental theorem. Then we define the product integral using a Riemann approach, prove the fundamental theorem, and give applications of the product integral. By following the standard approach in a typical first-year calculus sequence our hope is that the ways in which product calculus is like the usual calculus and the ways in which it is different will be more apparent. In fact, in many cases the product calculus version of a technique, result, or proof is identical to that of the corresponding technique, result, or proof for the usual calculus with the exception that addition, subtraction, multiplication, and division become multiplication, division, exponentiation, and taking roots, respectively.

Our treatment is not comprehensive; that would require something of book length. Our focus is on the basic ideas and applications of product calculus. In particular, we do just enough theory on the product derivative and integral to prove the fundamental theorem relating the two. (This includes product versions of many of the standard results in the first-year calculus sequence, though.) We take the Riemann approach to the integral, as is typical in a first calculus course, rather than the more sophisticated approaches of Lebesgue and others. We do not address multivariate functions or functions of a complex variable. We do not develop product versions of the standard integration techniques. While we prove a Taylor-like theorem for infinite products, we do not attempt to deal with the remainder function.

Definition and interpretation of the product derivative

The definition of the usual derivative function of f is

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

and the definition of the usual Riemann integral is

$$\int_a^b f(x)\, dx = \lim_{||P|| \to 0} \sum_{k=1}^{n} f(c_k) \Delta x_k.$$

There are several things that go into making the fundamental theorem of calculus work, but one of the most crucial is the relationship between the operations in the derivative and integral definitions. The derivative has subtraction where the integral has addition, and the derivative has division where the integral has multiplication. The importance of these operations appearing where they do is clear in certain proofs of the fundamental theorem of calculus. For example, the proof in the standard calculus text Strauss, Bradley, and Smith [8, p. 303] requires that a particular sum telescope, and precisely what makes the sum telescope is that addition, subtraction, multiplication, and division are in their appropriate places in the derivative and integral definitions.

With this in mind, then, if we want to attempt a different kind of calculus we should preserve the relationships between the operations in the derivative and integral definitions. We start with the derivative. Subtraction is to division what division is to taking roots, so let's change these operations accordingly in the definition of the usual derivative and define the result to be the *product derivative* function of f:

$$f^*(x) = dy^{1/dx} = \lim_{\Delta x \to 0} \left(\frac{f(x + \Delta x)}{f(x)} \right)^{\frac{1}{\Delta x}}.$$

(Of course, it's not clear yet why "product derivative" makes sense as a name for f^*, but that should become apparent shortly.) If $f^*(x)$ exists, then we say that f is *product-differentiable* at x. We also need a notation for successive product derivatives. Let $f^{**}(x)$ be the second product derivative of f; i.e., the product derivative of the product derivative of f. In general, let $f^{[n]}(x)$ denote the n^{th} product derivative of f at x. In addition, we can define right and left product derivatives by taking the limit as Δx approaches 0 from the right and from the left, respectively.

What does f^* mean? Well, as $f(x+\Delta x) - f(x)$ gives the change in f, in an additive sense, over the interval $[x, x + \Delta x]$,

$$\frac{f(x + \Delta x)}{f(x)}$$

gives the *multiplicative* change in f over $[x, x + \Delta x]$. For example, if $f(1) = 2$ and $f(3) = 18$, then the multiplicative change in f over $[1, 3]$ is 9, obtained by $f(3)/f(1)$. The function f changes by a factor of 9 as x changes from 1 to 3. Similarly, if $f(1) = 2$ and $f(3) = 1/2$, then the multiplicative change in f over $[1, 3]$ is $1/4$.

Then, just as

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

gives the average additive change in f per unit change in x over $[x, x + \Delta x]$,

$$\left(\frac{f(x + \Delta x)}{f(x)}\right)^{\frac{1}{\Delta x}}$$

gives the average multiplicative change in f per unit change of x over $[x, x + \Delta x]$. For example, if $f(1) = 2$ and $f(3) = 18$, then the average multiplicative change in f over $[1, 3]$ is $\sqrt{18/2} = 3$. On average, f changes by a factor of 3 as x changes from 1 to 2 and by another factor of 3 as x changes from 2 to 3.

If we take the limit as Δx approaches 0, then we obtain

$$f^*(x) = dy^{1/dx} = \lim_{\Delta x \to 0} \left(\frac{f(x + \Delta x)}{f(x)}\right)^{\frac{1}{\Delta x}},$$

which must be the instantaneous multiplicative change, or *instantaneous growth factor*, of f with respect to x. (If $f^*(x) < 1$, then f^* is, strictly speaking, a decay factor rather than a growth factor. However, for simplicity's sake we will refer to f^* as a growth factor, with the understanding that a growth factor smaller than one indicates decay, just as a negative increase indicates a decrease.)

It is also clear from the definition that $f^*(x)$ cannot exist if $f(x) = 0$. This makes sense with the interpretation of f^* as the instantaneous multiplicative change of f. However, $f^*(x)$ can exist if $f(x)$ is negative, provided that $f(x + \Delta x)$ is also negative for all sufficiently small values of Δx.

On the other hand, $f^*(x)$ itself cannot be negative.

Lemma 1. *If $f^*(c)$ exists, then $f^*(c) \geq 0$.*

Proof. Suppose $f^*(c)$ is negative. By the definition of limit there exists some $\delta > 0$ such that for all values of Δx with $0 < |\Delta x| < \delta$, $(f(c + \Delta x)/f(c))^{1/\Delta x} < 0$. However, for any $\delta > 0$, there exists an odd positive integer q such that $0 < \frac{2}{q} < \delta$. But $(f(c+2/q)/f(c))^{q/2}$ is the square root of some real number and thus cannot be negative. This is a contradiction, and thus $f^*(c)$ cannot be negative. □

Lemma 1 should make sense with our interpretation of f^* as a growth factor.

Rules for product differentiation

Like usual differentiation, there are some simple rules for calculating certain product derivatives. In this section we derive a few of these, and then we give a general way to determine a product derivative from the usual derivative.

The product derivative of a constant function is 1:

$$d(k)^{1/dx} = \lim_{\Delta x \to 0} \left(\frac{k}{k}\right)^{\frac{1}{\Delta x}} = \lim_{\Delta x \to 0} 1^{1/\Delta x} = 1.$$

Multiplying a function by a constant factor does not affect the product derivative:

$$d(kf(x))^{1/dx} = \lim_{\Delta x \to 0} \left(\frac{kf(x+\Delta x)}{kf(x)}\right)^{\frac{1}{\Delta x}} = d(f(x))^{1/dx}.$$

Exponential functions have constant product derivatives:

$$d(a^x)^{1/dx} = \lim_{\Delta x \to 0} \left(\frac{a^{x+\Delta x}}{a^x}\right)^{\frac{1}{\Delta x}} = \lim_{\Delta x \to 0} \left(a^{\Delta x}\right)^{\frac{1}{\Delta x}} = a.$$

The derivative of a product is now, as calculus students everywhere have wished, the product of the derivatives:

$$d(f(x)g(x))^{1/dx} = \lim_{\Delta x \to 0} \left(\frac{f(x+\Delta x)g(x+\Delta x)}{f(x)g(x)}\right)^{\frac{1}{\Delta x}} = d(f(x))^{1/dx} d(g(x))^{1/dx}.$$

All of these rules are intuitively consistent with our interpretation of $f^*(x)$ as the instantaneous growth factor of f.

Unfortunately, the product derivative of a linear function is a little more complicated, and our intuition starts to break down. For example, what is the instantaneous growth factor of the function $f(x) = x$? Applying the definition to obtain the product derivative requires previous knowledge of the value of a specific limit expression. We have

$$d(x)^{1/dx} = \lim_{\Delta x \to 0} \left(\frac{x+\Delta x}{x}\right)^{\frac{1}{\Delta x}} = \lim_{\Delta x \to 0} \left(1 + \frac{\Delta x}{x}\right)^{\frac{1}{\Delta x}} = e^{1/x}.$$

We could prove some other product differentiation rules directly, but many of the derivations, like that of $f(x) = x$, would require dealing with the indeterminate form 1^∞. Rather than dealing with all of these indeterminate forms separately, there is a simpler approach, based on the following relationship between f^* and the usual derivative f'.

Theorem 1. *If $f(x) \neq 0$, then $f^*(x)$ exists and is nonzero if and only if $f'(x)$ exists, in which case*

$$f^*(x) = \exp\left(\frac{d}{dx} \ln|f(x)|\right) = e^{f'(x)/f(x)}.$$

Proof. Suppose $f^*(x)$ exists and is nonzero. An argument similar to that in the proof of Lemma 1 shows that $f(x+\Delta x)$ and $f(x)$ must have the same sign for all sufficiently small Δx. Thus we have

$$\begin{aligned}
\ln f^*(x) &= \ln\left(\lim_{\Delta x\to 0}\left(\frac{f(x+\Delta x)}{f(x)}\right)^{\frac{1}{\Delta x}}\right)\\
&= \lim_{\Delta x\to 0}\ln\left(\frac{|f(x+\Delta x)|}{|f(x)|}\right)^{\frac{1}{\Delta x}}\\
&= \lim_{\Delta x\to 0}\left(\frac{\ln|f(x+\Delta x)|-\ln|f(x)|}{\Delta x}\right)\\
&= \frac{d}{dx}\left(\ln|f(x)|\right).
\end{aligned}$$

(We can interchange ln and lim in the second step because $\ln x$ is continuous.) Since $\ln|f(x)|$ is differentiable, the chain rule implies that $|f(x)| = e^{\ln|f(x)|}$ must also be differentiable. As $f(x)\neq 0$, the differentiability of $|f|$ at x implies that of f at x. Differentiating $\ln|f(x)|$, then, we obtain

$$\ln f^*(x) = \frac{f'(x)}{f(x)},$$

which implies

$$f^*(x) = e^{f'(x)/f(x)}.$$

A straightforward reversal of these steps shows that the existence of $f'(x)$ implies the existence of $f^*(x)$ and that $f^*(x) = e^{f'(x)/f(x)}$. Since $e^{f'(x)/f(x)}$ can never be 0, the existence of $f'(x)$ also implies that $f^*(x)$ is nonzero. □

With Theorem 1, we can establish other product differentiation rules. For example, here is the sum rule for product differentiation:

$$d(f(x)+g(x))^{1/dx} = \exp\left(\frac{f'(x)+g'(x)}{f(x)+g(x)}\right).$$

The power rule for product differentiation is

$$d(x^n)^{1/dx} = e^{n/x}.$$

We also have product differentiation rules for sine and cosine:

$$d(\sin x)^{1/dx} = e^{\cot x},\ \text{and}$$

$$d(\cos x)^{1/dx} = e^{-\tan x}.$$

The function e^{e^x}, rather than e^x, is its own product derivative:

$$d(e^{e^x})^{1/dx} = \exp\left(e^x(e^{e^x})/e^{e^x}\right) = e^{e^x}.$$

The function x^x has an interesting product derivative as well:

$$d(x^x)^{1/dx} = \exp\left(\frac{d}{dx}(x \ln x)\right) = e^{\ln x + 1} = ex.$$

This allows us to determine the function whose product derivative is x, namely, $f(x) = (x/e)^x$.

$$d\left(\left(\frac{x}{e}\right)^x\right)^{1/dx} = \frac{d(x^x)^{dx}}{d(e^x)^{dx}} = \frac{ex}{e} = x.$$

Higher-order derivatives may also be calculated via Theorem 1.

Corollary 1. *If $f(x) \neq 0$, then $f^{[n]}(x)$ exists and is nonzero if and only if $f^{(n)}(x)$ exists, in which case*

$$f^{[n]}(x) = \exp\left(\frac{d^n}{dx^n}(\ln|f(x)|)\right).$$

Proof. Theorem 1 covers the case $n = 1$. Since $f^*(x) = e^{f'(x)/f(x)}$ cannot be 0, we can invoke Theorem 1 to prove the case $n = 2$. Induction covers the remaining cases. □

Applications of the product derivative

Is the product derivative good for anything? We discuss three applications and briefly touch on a fourth.

Maximizing investment profit

One application involves investing. If one has a fixed amount of money to invest in some commodity, when should that money be invested to maximize profit? The time at which the price of the commodity is lowest is not necessarily optimal, as profit comes from growth, not simply from "buying low." However, when the amount of money is fixed the greatest absolute growth does not necessarily yield the maximum profit, either. For example, spending \$100 to purchase a commodity at \$10 a share and selling it at \$50 a share produces a much larger profit, \$400, than buying the commodity at \$100 and selling it at \$200, which only produces a profit of \$100. The reason, of course, is that the higher price in the latter case means that one cannot buy nearly as much of the commodity, offsetting the greater absolute price increase. As we can probably see from this example, maximum profit in the presence of a fixed amount to invest actually occurs when the multiplicative growth is greatest. This relates to the product derivative, as the greatest instantaneous multiplicative growth occurs at the value of t for which $f^*(t)$ is maximized. Of course, in real situations investing will be done over a time period; thus to determine maximum profit one must find a time interval, not an instant, over which multiplicative growth is greatest. However, if the length of the time interval is small then the product derivative can be used to give a good first approximation, as the optimal time interval probably has the time at which $f^*(t)$ is maximized somewhere near its midpoint.

For example, suppose the unit price of some commodity as a function of time has been predicted to behave as $f(t) = t^4 - 4t^3 + 50$, and you have \$100 to invest. If you have decided that you want your money tied up in the investment for 0.5 units of time, at which time should you buy the commodity? Simply by looking at a graph of f, as in Figure 1, it is difficult to determine where the multiplicative growth is greatest. We

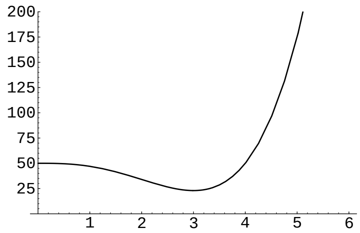

Figure 1: $f(t) = t^4 - 4t^3 + 50$

can calculate the product derivative by Theorem 1, though, obtaining

$$f^*(t) = \exp\left(\frac{4t^3 - 12t^2}{t^4 - 4t^3 + 50}\right).$$

Maximizing this expression is a bit nasty algebraically, but it's easy to get a good numerical approximation using technology; the maximum value occurs around $t = 4.2$. Thus we should probably buy sometime around the time $t = 4$ in order to maximize profit. At this time, the price is \$50, so we can buy 100/50, or 2, units. At time $t = 4.5$, the price is \$95.5625. So we make $100/50(95.5625) - 100$, or (to the nearest cent) \$91.13. This is a much greater profit than purchasing when the price is lowest, at time $t = 3$; the profit then is only \$24.18. And buying several time periods later at $t = 10$, when the absolute growth is much greater, produces a profit of only \$25.20. A more detailed analysis without using the product derivative could, of course, improve on the profit of \$91.13; the point is that we get a good initial estimate by using the product derivative. (In fact, in this example our guess was quite good; the optimal time interval buys at time 3.98, producing a profit of \$91.14.)

In general, the smaller the investment time window, the closer the optimal investment time comes to the time at which the product derivative is maximized. (It is also worth noting that maximizing $f^*(t) = \exp(f'(t)/f(t))$ is equivalent to maximizing $f'(t)/f(t)$, the instantaneous percentage growth of f, as e^x is an increasing function of x.)

Fitting exponential functions to curves and a product version of Taylor series

Another application of the product derivative involves approximating the graphs of functions by exponential curves, just as the usual derivative can be used to help approximate the graphs of functions by straight lines. To take a simple example, let's find the best exponential function approximation $g_1(x)$ to $f(x) = x$ when $x = 1$, where "best" means that, at $x = 1$, g and f share the same function value and the same instantaneous growth factor. Let $g_1(x) = ka^x$ be a generic exponential function. The problem then becomes determining the values of k and a such that $g_1(1) = f(1)$ and $g_1^*(1) = f^*(1)$. We have $f^*(x) = e^{1/x}$, with $f^*(1) = e$. Thus $g_1(1) = 1$ and $g_1^*(1) = e$. Since $g_1^*(x) = a$, this yields $a = e$ and $k = 1/e$. Therefore, $g_1(x) = e^{x-1}$. A graph of f and g_1 is given in Figure 2. As we can see, g_1 does closely approximate f near 1.

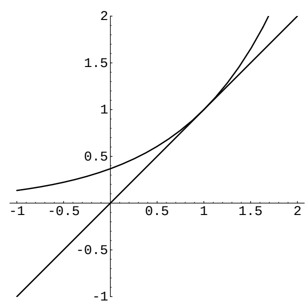

Figure 2: First-order exponential approximation of $f(x) = x$

It's not too hard to see how to generalize what we've just done: The best exponential approximation g to the function f near c (where "best" is defined as above) is given by $g_1(x) = f(c)f^*(c)^{x-c}$. This expression, of course, is similar to that for the best linear approximation L to f near c: $L(x) = f(c) + f'(c)(x - c)$. In fact, changing addition to multiplication, multiplication to exponentiation, and the usual derivative to the product derivative in the latter expression produces the former.

The next step after approximating a function f with a linear function via the usual derivative is approximating f with a quadratic, a cubic, etc., via higher-order derivatives. This leads to the Taylor series representation of f. We can obtain a Taylor series-like representation of f by taking higher-order product derivatives as well. Following the pattern in the previous example, the proper form for g_2 in order for f, f^*, and f^{**} to agree with g_2, g_2^*, and g_2^{**} at c, respectively, is $g_2(x) = f(c)f^*(c)^{x-c}f^{**}(c)^{1/2(x-c)^2}$.

With $f(x) = x$, we have

$$f^{**}(x) = \exp\left(\frac{d}{dx}\ln e^{1/x}\right) = e^{-1/x^2},$$

yielding $f^{**}(1) = e^{-1}$ and $g_2(x) = e^{(x-1)-1/2(x-1)^2}$. A graph of f and g_2 are given in Figure 3. As we can see, including the second-order information from f^{**} produces a

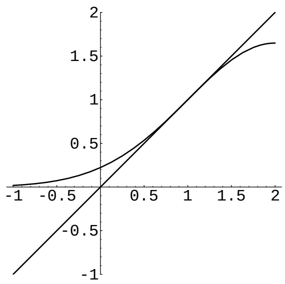

Figure 3: Second-order exponential approximation of $f(x) = x$

better approximation of f near $x = 1$.

It's not too hard to observe at this point that the best third-order approximation is $g_3(x) = f(c)f^*(c)^{x-c}f^{**}(c)^{1/2(x-c)^2}f^{***}(c)^{(x-c)^3/3!}$. Generalizing, the best order n approximation is given by

$$g_n(x) = \prod_{i=0}^{n}\left(f^{[i]}(c)\right)^{(x-c)^i/i!}.$$

This equation is the direct product analogue of the order n Taylor approximation of f, where addition, multiplication, and the derivative have been replaced by multiplication, exponentiation, and the product derivative, respectively.

However, $g_n(x)$ can also be stated explicitly in terms of the usual Taylor approximation of $\ln|f(x)|$.

Theorem 2. *Let f be n-times differentiable and nonzero at c and let $T_n(x)$ be the usual n^{th}-order Taylor approximation of $h(x) = \ln|f(x)|$ about c. Then the best n^{th}-order exponential function approximation g_n to f near c can be expressed as $g_n(x) = e^{T_n(x)}$.*

Proof.

$$\begin{aligned} g_n(x) &= \prod_{i=0}^{n}\left(f^{[i]}(c)\right)^{(x-c)^i/i!} = \prod_{i=0}^{n}\left(\exp(h^{(i)}(c))\right)^{(x-c)^i/i!} = \exp\sum_{i=0}^{n}\frac{h^{(i)}(c)}{i!}(x-c)^i \\ &= e^{T_n(x)}. \end{aligned}$$

□

For example, the second-order exponential approximation of $f(x) = x$ near $x = 1$, $g_2(x) = e^{(x-1)-1/2(x-1)^2}$, has the second-order Taylor series expansion of $\ln x$ about $x = 1$ in the exponent.

We can also use Theorem 2 to obtain higher-order exponential approximations of $f(x) = x$ quite quickly. For example, about $x = 1$ we have

$$g_{10}(x) = \exp\left(\sum_{i=1}^{10}(-1)^{i-1}\frac{(x-1)^i}{i}\right).$$

The graphs of f and g_{10} are given in Figure 4. It appears we are obtaining convergence

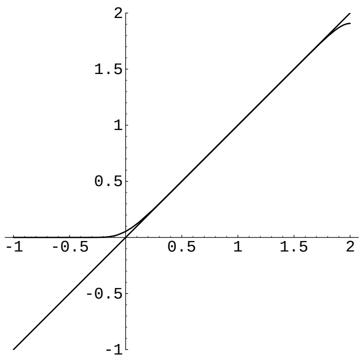

Figure 4: Tenth-order exponential approximation of $f(x) = x$

on the interval (0,2) and possibly at 0 and 2. Since (0,2] is the interval of convergence of $\ln x$ about 1, this is to be expected.

Slopes on semilog graphs

A third application of the product derivative involves log-linear, or semilog, graphs. A *semilog graph* is a graph in which the horizontal scale is linear and the vertical scale is logarithmic. This means that two points equally distant on the horizontal scale have differences in the horizontal variable that are the same, while two points that are equally distant on the vertical scale have ratios in the vertical variable that are the same. From a geometric standpoint, a plot of f on a semilog graph looks like a plot of $\log f$ on a usual linear-linear graph. For example, Figure 5 is a semilog plot of the function $f(x) = x$. Geometrically, it looks like the graph of $\log x$.

Probably the most important aspect of semilog plots is that exponential functions appear as straight lines, with slopes proportional to the exponential base. For example, Figure 6 contains a semilog plot of $f(x) = 2^x$ and $g(x) = 3^x$. The latter has the steeper slope.

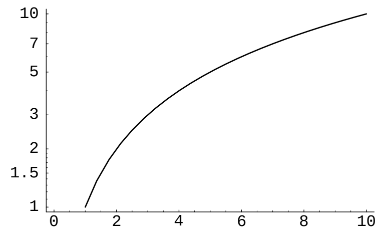

Figure 5: Semilog plot of $f(x) = x$

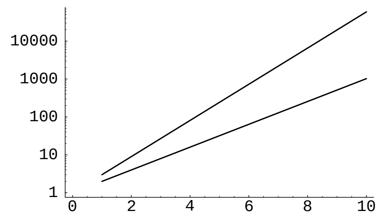

Figure 6: Semilog plots of 2^x and 3^x

Experimental scientists use semilog plots to discover or verify that a set of collected data has an exponential relationship. The difference between a set of data with an exponential relationship and one with a quadratic or higher polynomial relationship can be difficult to distinguish on a linear-linear plot. However, plotting the data on a semilog graph produces a straight line if the data are exponential and something nonlinear if the data are polynomial. (For just one example of many, see Wang, Molina, and Abdalla [10].)

The product derivative's connection to semilog graphs is that, in a certain sense, $f^*(c)$ gives the slope of the graph of f at c when f is plotted on a semilog graph, provided $f^*(c) \neq 0$. We say "in a certain sense" because there are some scaling issues that we have to be careful with. We will deal with these by defining slope and then function value in a purely geometric sense.

The usual definition of slope is independent of how f is graphed. This is as it should be, as mathematical chaos would insue if the slope of a function were allowed to change depending on the scaling of the axes we use in drawing it. On the other hand, this definition has the drawback that if the horizontal and vertical scaling are very different, then what the slope of f looks like from the graph can be quite different from what it actually is. We can see this in Figure 5. The slope of the curve is 1 everywhere, but it doesn't look like it because of the logarithmic scaling on the vertical axis.

It's kind of odd, though, to say that something "doesn't look like" what it actually is. So what are we really doing when we look at a graph like Figure 5 and saying that the slope "doesn't look like" it's 1? We're superimposing on the graph a linear scale identical in both horizontal and vertical directions, and we're reading the graph with respect to that linear scale.

This observation is the motivation for our definition of geometric slope. Given a function f and specific horizontal and vertical scales, draw f with respect to these scales. Then superimpose a linear scale identical in both horizontal and vertical directions on this graph of f. The *geometric slope* of f at a point is then the slope of f at that point measured with respect to the superimposed linear scale. The geometric slope of a curve thus corresponds to what the slope appears to be when one looks at the curve and ignores the scales that were used to plot it.

Function value is also independent of the scales used to plot a function f. For our purposes, though, we need a concept to represent the apparent height of the graph of a function when it is drawn with respect to certain horizontal and vertical scales. Our definition of this geometric value takes the same approach as that for geometric slope. However, unlike geometric slope, geometric value is dependent on the specific superimposed linear scale. (Geometric slope has a dimension cancellation in its calculation and thus is the same regardless of the choice of linear scale.) We define the *geometric value* of f at a point with a given horizontal and vertical scale and a given superimposed linear scale identical in both directions to be the height of that point when f is plotted using the given horizontal and vertical scale but then measured with respect to the linear scale.

With geometric slope and value defined we can now state the connection between the product derivative and semilog graphs. *If $f^*(c) \neq 0$, then the geometric value of $f^*(c)$ is the geometric slope of the graph of f at c when f is plotted on a semilog graph.* The specific linear scale required for geometric value is the same as that of the horizontal scale from the semilog graph.

To see this, let's start with the semilog plot of $f(x) = x$ and its product derivative $f^*(x) = e^{1/x}$ in Figure 7. Considering the graphs of f and f^* purely geometrically, the

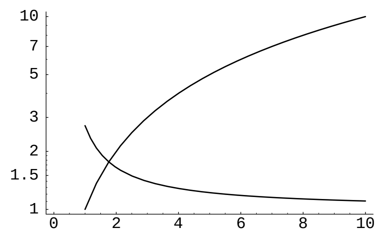

Figure 7: Semilog plots of x and $e^{1/x}$

graph of f^* certainly behaves like the graph of a function that would give the slopes of f. It takes its largest values when the graph of f is steepest, it decreases as the slopes of f decrease, and it's positive everywhere, corresponding with the fact that f is increasing.

We can see this more precisely via Theorem 1. Geometrically, the graph of f on a semilog plot with e as the base of the logarithm looks identical to the graph of $\ln f$ on a linear-linear plot. Thus the geometric slopes of the graph of f on a semilog plot are the same as those of $\ln f$ on a linear-linear plot. Therefore, $\frac{d}{dx}(\ln f(x))$ gives the geometric slopes of the graph of f on a semilog plot. In order to obtain a graph that looks like $\frac{d}{dx}(\ln f(x))$ when plotted on the same semilog graph as f, though; i.e., one whose geometric values are these slopes, we need to scale $\frac{d}{dx}(\ln f(x))$ exponentially. Doing so produces $\exp(\frac{d}{dx}(\ln f(x)))$, which, according to Theorem 1, is $f^*(x)$ (subject, of course, to the restrictions given in the statement of the theorem).

Even though the previous argument assumes that e is the base of the logarithm, the interpretation of the geometric value of the product derivative as the geometric slope of the graph of f on a semilog plot actually holds regardless of the base used. Suppose we use b as the logarithm base on a semilog plot of f. Then the geometric slope of f is

$$\frac{d}{dx}(\log_b f(x)) = \frac{f'(x)}{f(x)\ln b}.$$

A plot of f^* on the same semilog graph will have geometric value

$$\log_b\left(e^{f'(x)/f(x)}\right) = \frac{f'(x)}{f(x)}\log_b e = \frac{f'(x)}{f(x)}\frac{\ln e}{\ln b} = \frac{f'(x)}{f(x)\ln b},$$

the same as the geometric slope of f.

With this interpretation of the product derivative we can attempt a more general view of the relationship between derivatives and slopes. The geometric values of $f'(c)$ and $f^*(c)$ both give the geometric slope of f at c. However, with f', the original scaling used to draw the graph of f is linear-linear, and with f^*, the original scaling is semilog or log-linear.

Optimization

Although there is no real advantage over the approach with usual differentiation, one can also use the product derivative to do optimization. We have

Theorem 3. *If f has a local maximum or local minimum at an interior point c of its domain and $f^*(c)$ exists, then $f^*(c) = 1$.*

Proof. Since f^* is product-differentiable at c, $f(c) \neq 0$. Suppose f has a local maximum at c. Then $f(c) \geq f(c+\Delta x)$ and $f(c)$ and $f(c+\Delta x)$ have the same sign for all sufficiently small Δx. Fix such a Δx. Then

$$\frac{f(c+\Delta x)}{f(c)} \leq 1 \Rightarrow \left(\frac{f(c+\Delta x)}{f(c)}\right)^{\frac{1}{\Delta x}} \leq 1^{1/\Delta x} = 1,$$

if Δx is positive. Taking the limit, we obtain

$$\lim_{\Delta x \to 0^+} \left(\frac{f(c+\Delta x)}{f(c)} \right)^{\frac{1}{\Delta x}} \leq 1 \Rightarrow f^*(c) \leq 1.$$

A similar argument shows that if Δx is negative, then $f^*(c) \geq 1$. Thus $f^*(c) = 1$.

If f has a local minimum at c, then a similar argument shows that $f^*(c) = 1$ as well. □

By Theorem 1, solving $f^*(x) = 1$ for $f(x) \neq 0$ is equivalent to solving $e^{f'(x)/f(x)} = 1$, which, after taking logs and multiplying by $f(x)$, reduces to the problem of solving $f'(x) = 0$.

Some theory on the product derivative

We need to establish a few results about the product derivative in order to prove part of the fundamental theorem of product calculus. These are product derivative versions of most of the standard results on the usual derivative discussed in a first-semester calculus course. The proofs are virtually identical to the proofs of the corresponding results on the usual derivative, too. Almost all of the changes are the obvious ones that would need to be made as addition, subtraction, multiplication, and division become multiplication, division, exponentiation, and taking roots, respectively. Alternatively, we could use the relationship between the product derivative and the usual derivative given in Theorem 1 to obtain these results. However, it is worth seeing explicitly that the results for the product derivative parallel, but do not depend on, the corresponding results for the usual derivative.

Theorem 4. *(Product differentiability and continuity.) If f is product-differentiable at c and $f^*(c) > 0$, then f is continuous at c.*

Proof. Since f is product-differentiable at c, $f(c) \neq 0$. Then

$$\begin{aligned} f(c+\Delta x) &= f(c)\frac{f(c+\Delta x)}{f(c)} \\ &= f(c)\left(\left(\frac{f(c+\Delta x)}{f(c)}\right)^{\frac{1}{\Delta x}}\right)^{\Delta x}. \end{aligned}$$

Taking the limit as Δx approaches 0, we have

$$\lim_{\Delta x \to 0} f(c+\Delta x) = \lim_{\Delta x \to 0} f(c)\left(\left(\frac{f(c+\Delta x)}{f(c)}\right)^{\frac{1}{\Delta x}}\right)^{\Delta x}$$

$$\begin{aligned}&= f(c) \lim_{\Delta x \to 0} \left(\left(\frac{f(c+\Delta x)}{f(c)} \right)^{\frac{1}{\Delta x}} \right)^{\Delta x} \\ &= f(c) f^*(c)^0 \\ &= f(c).\end{aligned}$$

By definition, then, f is continuous at c.

Similar arguments, using one-sided limits, show that product-differentiability of f from the right and left imply right- and left-continuity, respectively, of f at c. □

The requirement in Theorem 4 that $f^*(c)$ be positive is necessary, as it is possible for f^* to be 0 at points where f is not continuous. For example, let

$$f(x) = \begin{cases} 3, & x < 1; \\ 2, & x = 1; \\ 1, & x > 1. \end{cases}$$

Then

$$\lim_{\Delta x \to 0^-} \left(\frac{f(1+\Delta x)}{f(1)} \right)^{\frac{1}{\Delta x}} = \lim_{\Delta x \to 0^-} \left(\frac{3}{2} \right)^{\frac{1}{\Delta x}} = \lim_{\Delta x \to 0^+} \left(\frac{2}{3} \right)^{\frac{1}{\Delta x}} = 0.$$

Similarly,

$$\lim_{\Delta x \to 0^+} \left(\frac{f(1+\Delta x)}{f(1)} \right)^{\frac{1}{\Delta x}} = \lim_{\Delta x \to 0^+} \left(\frac{1}{2} \right)^{\frac{1}{\Delta x}} = 0.$$

Thus $f^*(1) = 0$. In fact, any type of jump discontinuity like that exhibited by f at 1 will produce a product derivative of 0.

Theorem 5. *(Rolle's Theorem for Product Derivatives.) If f is continuous on $[a, b]$, product-differentiable on (a, b), and $f(a) = f(b)$, then there exists some c in (a, b) such that $f^*(c) = 1$.*

As usual, Rolle's Theorem says that continuity of a function on an interval for which the values at the endpoints are the same means that the function must "turn around" somewhere in the interval. "Turning around" means a local maximum or minimum, which in the product derivative case yields an f^* value of 1 rather than the 0 that occurs with the usual derivative.

Proof. By the Extreme Value Theorem, f has a maximum and a minimum on $[a, b]$. By Theorem 3, these occur either at 1) a and b, 2) points on the interior of (a, b) where f^* is undefined, or 3) points where $f^* = 1$. Case 3 satisfies the theorem. We can rule out Case 2 because f is product-differentiable on (a, b). If the maximum and minimum both occur at the endpoints, then one of a or b is the maximum, and the other is the minimum. But $f(a) = f(b)$, so this means that f must be a constant function on (a, b). Then we must have $f^*(x) = 1$ for every point in (a, b), also satisfying the theorem. □

Theorem 6. *(Mean Value Theorem for Product Derivatives.) If f is continuous on $[a,b]$, product-differentiable on (a,b), $f(a) \neq 0$, and $f(b) \neq 0$, then there exists some c in (a,b) such that*

$$f^*(c) = \left(\frac{f(b)}{f(a)}\right)^{1/(b-a)}.$$

The expression on the right is the average multiplicative change of f over $[a,b]$. Thus the theorem states that at some point in (a,b) the product derivative attains the average multiplicative change, just as under corresponding conditions the usual Mean Value Theorem states that the usual derivative attains the average additive change.

Proof. Let

$$h(x) = \frac{f(x)}{f(a)}\left(\frac{f(a)}{f(b)}\right)^{(x-a)/(b-a)}.$$

The function h is continuous on $[a,b]$ because it is composed of continuous functions, and it is product-differentiable on (a,b) for a similar reason. Moreover, $h(a) = h(b) = 1$. The hypotheses of Rolle's Theorem for Product Derivatives are thus satisfied, and so there exists some c in (a,b) such that $h^*(c) = 1$. Product-differentiating the expression on the right, we obtain

$$h^*(x) = f^*(x)\left(\frac{f(a)}{f(b)}\right)^{1/(b-a)}.$$

Substituting c and using the fact that $h^*(c) = 1$ produces

$$1 = f^*(c)\left(\frac{f(a)}{f(b)}\right)^{1/(b-a)},$$

which then implies

$$f^*(c) = \left(\frac{f(b)}{f(a)}\right)^{1/(b-a)}.$$

□

As with the usual Mean Value Theorem, the Mean Value Theorem for Product Derivatives yields two quick corollaries.

Corollary 2. *If $f^*(x) = 1$ for all x in (a,b), then there exists some constant C such that $f(x) = C$ for all x in (a,b).*

Proof. Let x_1 and x_2 be any two points in (a,b) with $x_1 < x_2$. By Theorem 4, f is continuous on $[x_1, x_2]$. Also, since f^* exists at x_1 and x_2, f is nonzero at x_1 and at

x_2. The function f therefore satisfies the hypotheses of the Mean Value Theorem for Product Derivatives on $[x_1, x_2]$, which implies that there exists c in (x_1, x_2) such that

$$f^*(c) = \left(\frac{f(x_2)}{f(x_1)}\right)^{1/(x_2-x_1)}.$$

But $f^*(c) = 1$, by hypothesis. Thus we must have $f(x_1) = f(x_2)$.

Since this argument may be made for any points x_1 and x_2 in (a, b), we must have $f(x_1) = f(x_2)$ for all x_1, x_2 in (a, b). Therefore, f is constant on (a, b). □

Corollary 3. *If $f^*(x) = g^*(x) \neq 0$ for all x in (a, b), then there exists some constant K such that $f(x) = Kg(x)$ for all x in (a, b).*

This result says that if two functions have the same product derivative, then the ratio of the two functions is constant.

Proof. Since $g^*(x)$ exists for all x in (a, b), $g(x) \neq 0$ for any x in (a, b). Let $h(x) = f(x)/g(x)$. Then $h^*(x) = f^*(x)/g^*(x) = 1$ for all x in (a, b). By Corollary 2, there exists some constant K such that $h(x) = K$ on (a, b). Thus, on (a, b), $f(x)/g(x) = K$, and $f(x) = Kg(x)$. □

Definition of the product integral

We can also define a product integral in an fashion analogous to the way in which we defined the product derivative. The definition of the Riemann integral of f over $[a, b]$ is the following: Let $P = \{x_0, x_1, \ldots, x_n\}$ be a partition of $[a, b]$, and let c_k be any point in $[x_{k-1}, x_k]$. Let Δx_k be $x_k - x_{k-1}$, the width of subinterval k in P. Let $||P||$ be the maximum value of Δx_k; i.e., the width of the largest subinterval in P. Then the Riemann integral of f over $[a, b]$ is given by

$$\int_a^b f(x)\,dx = \lim_{||P|| \to 0} \sum_{k=1}^{n} f(c_k)\Delta x_k,$$

provided the limit exists and is independent of the choice of P and the c_k's.

Just as we changed subtraction to division and division to taking roots in the definition of the usual derivative in order to define the product derivative, we change addition to multiplication and multiplication to exponentiation to define the product integral. This yields the *Riemann product integral* of f over $[a, b]$:

$$\mathscr{P}_a^b f(x)^{dx} = \lim_{||P|| \to 0} \prod_{k=1}^{n} f(c_k)^{\Delta x_k},$$

provided the limit exists and is independent of the choice of P and the c_k's. (See Gill and Johansen [3] and Guenther [5] for another way to define the product integral, although the latter shows how to convert between the two types.) This definition also implies that f must be non-negative over $[a, b]$ in order for $\mathscr{P}_a^b f(x)^{dx}$ to exist, for the same reason as in the proof of Lemma 1. Moreover, if $f(x) = 0$ for a finite number of values of x, then it is possible to choose partitions of $[a, b]$ and corresponding $f(c_k)$ values that do and that do not include the zeros. Those that do include the zeros will yield a limiting value of 0, and those that do not may not. Thus if $f(x) = 0$ for a finite number of values of x, then $\mathscr{P}_a^b f(x)^{dx}$ may not exist. We will preclude this possibility, then, by considering functions f that only take on positive values on $[a, b]$. Thus if $f(x) > 0$ on $[a, b]$ and $\mathscr{P}_a^b f(x)^{dx}$ exists, then we say f is *product-integrable* over $[a, b]$.

What does $\mathscr{P}_a^b f(x)^{dx}$ mean, though? A somewhat general view of $\int_a^b f(x)\,dx$ requires thinking of f as the additive rate of change of some quantity Q with respect to x. Then $\int_a^b f(x)\,dx$ represents the total accumulation of Q as x ranges from a to b. With this in mind, we can think of f as the multiplicative rate of change of some quantity Q with respect to x. Then $f(c_k)^{\Delta x_k}$ gives an approximation of the multiplicative accumulation of Q over subinterval k. For example, if $f(x) = 3$ and $\Delta x_k = 2$, then the total multiplicative accumulation of Q over subinterval k is exactly $3(3) = 3^2 = 9$. The expression $\prod_{k=1}^n f(c_k)^{\Delta x_k}$ then gives an approximation of the multiplicative accumulation of Q over the entire interval $[a, b]$. Taking the limit as the width of the largest subinterval goes to 0, obtaining $\mathscr{P}_a^b f(x)^{dx}$, must then yield the exact value of the total multiplicative accumulation of Q over $[a, b]$.

A fundamental theorem for product calculus

Before we do any applications of the product integral, let's show that it is, in fact, the right integral for the product derivative, in the sense that there is a fundamental theorem relating the two. The proof is a straightforward modification of the proof of the fundamental theorem for the usual calculus in a standard calculus text [11], where, once again, addition, subtraction, multiplication, and division are replaced, respectively, by multiplication, division, exponentiation, and taking roots. An alternative way to proceed would be to establish an integral version of Theorem 1 first and then use appropriate properties of the usual derivative and integral. We intentionally avoid this route because, as with the results on the product derivative, we wish to show that the basic results of the product integral closely parallel but are independent of the usual integral.

First, a couple of definitions:

$$\mathscr{P}_a^a f(x)^{dx} = 1, \text{ and}$$
$$\mathscr{P}_b^a f(x)^{dx} = \frac{1}{\mathscr{P}_a^b f(x)^{dx}}, \text{ for } b > a.$$

Lemma 2. *(Comparison of product integrals.) If f and g are product-integrable on $[a, b]$ and $f(x) \leq g(x)$ for each x in $[a, b]$, then $\mathscr{P}_a^b f(x)^{dx} \leq \mathscr{P}_a^b g(x)^{dx}$.*

Proof. Let P be a partition of $[a, b]$, and let c_k be the representative from subinterval k. Then, by hypotheses, $f(c_k) \leq g(c_k)$ for each k. Thus

$$\begin{aligned} & \prod_{k=1}^{n} f(c_k)^{\Delta x_k} \leq \prod_{k=1}^{n} g(c_k)^{\Delta x_k} \\ \Rightarrow \quad & \lim_{||P|| \to 0} \prod_{k=1}^{n} f(c_k)^{\Delta x_k} \leq \lim_{||P|| \to 0} \prod_{k=1}^{n} g(c_k)^{\Delta x_k} \\ \Rightarrow \quad & \mathscr{P}_a^b f(x)^{dx} \leq \mathscr{P}_a^b g(x)^{dx}. \end{aligned}$$

□

Lemma 3. *(Product integrals of constant functions). If there exists some positive constant C such that $f(x) = C$ for all x in $[a, b]$, then $\mathscr{P}_a^b C^{dx} = C^{b-a}$.*

Proof. For any partition P and choice of c_k's, we have

$$\begin{aligned} \mathscr{P}_a^b f(x)^{dx} &= \lim_{||P|| \to 0} \prod_{k=1}^{n} f(c_k)^{\Delta x_k} = \lim_{||P|| \to 0} \prod_{k=1}^{n} C^{\Delta x_k} = \lim_{||P|| \to 0} C^{\sum_{k=1}^{n} \Delta x_k} \\ &= \lim_{||P|| \to 0} C^{b-a} = C^{b-a}. \end{aligned}$$

□

Lemma 4. *(Positivity of product integrals.) If f is continuous and product-integrable on $[a, b]$, then $\mathscr{P}_a^b f(x)^{dx} > 0$.*

Proof. The continuity of f, via the Extreme Value Theorem, implies that f attains its minimum on $[a, b]$. Since f is product-integrable on $[a, b]$, $f(x) > 0$ for all x in $[a, b]$. Thus there exists some constant $K > 0$ such that $f(x) \geq K$ for all x in $[a, b]$. By the two previous results, we have $\mathscr{P}_a^b f(x)^{dx} \geq K^{b-a} > 0$. □

Lemma 5. *(Multiplication of product integrals.) If f is product-integrable on $[a, c]$, then, for $a \leq b \leq c$,*

$$\mathscr{P}_a^c f(x)^{dx} = \left(\mathscr{P}_a^b f(x)^{dx}\right)\left(\mathscr{P}_b^c f(x)^{dx}\right).$$

Proof. Let $P_1 = \{x_0, x_1, \ldots, x_m\}$ and $P_2 = \{x_m, x_{m+1}, \ldots, x_n\}$ be partitions of $[a, b]$ and $[b, c]$, respectively. Then $P = P_1 \cup P_2$ is a partition of $[a, b]$. We have

$$\begin{aligned} \left(\mathscr{P}_a^b f(x)^{dx}\right)\left(\mathscr{P}_b^c f(x)^{dx}\right) &= \lim_{||P_1|| \to 0} \prod_{k=1}^{m} f(c_k)^{\Delta x_k} \lim_{||P_2|| \to 0} \prod_{k=m+1}^{n} f(c_k)^{\Delta x_k} \\ &= \lim_{\max\{||P_1||, ||P_2||\} \to 0} \prod_{k=1}^{n} f(c_k)^{\Delta x_k} \end{aligned}$$

$$
\begin{aligned}
&= \lim_{||P|| \to 0} \prod_{k=1}^{n} f(c_k)^{\Delta x_k} \\
&= \mathscr{P}_a^c f(x)^{dx}.
\end{aligned}
$$

□

Lemma 6. *(Max-Min Inequality.) If f is product-integrable on $[a, b]$ and has maximum $\max f$ and minimum $\min f$ on $[a, b]$, then*

$$(\min f)^{b-a} \leq \mathscr{P}_a^b f(x)^{dx} \leq (\max f)^{b-a}.$$

Proof. Let P be a partition of $[a, b]$, and let c_k be the representative point in partition k. Then

$$
\begin{aligned}
(\min f)^{b-a} &= (\min f)^{\sum_{k=1}^{n} \Delta x_k} = \prod_{k=1}^{n} (\min f)^{\Delta x_k} \leq \prod_{k=1}^{n} f(c_k)^{\Delta x_k} \leq \prod_{k=1}^{n} (\max f)^{\Delta x_k} \\
&= (\max f)^{\sum_{k=1}^{n} \Delta x_k} = (\max f)^{b-a}.
\end{aligned}
$$

Since the argument holds for any partition P and choice of c_k's, by the Squeeze Theorem we must have

$$(\min f)^{b-a} \leq \mathscr{P}_a^b f(x)^{dx} \leq (\max f)^{b-a}.$$

□

Theorem 7. *(Mean Value Theorem for Product Integrals.) If f is continuous and positive on $[a, b]$, then at some point c in $[a, b]$,*

$$f(c) = \left(\mathscr{P}_a^b f(x)^{dx}\right)^{1/(b-a)}.$$

Proof. Taking all three expressions in the Max-Min Inequality to the $1/(b-a)$ power, we obtain

$$\min f \leq \left(\mathscr{P}_a^b f(x)^{dx}\right)^{1/(b-a)} \leq \max f.$$

Since f is continuous on $[a, b]$, then, by the Intermediate Value Theorem, it must attain every value between $\min f$ and $\max f$, inclusive. In particular, then, there exists some c in $[a, b]$ such that $f(c) = \left(\mathscr{P}_a^b f(x)^{dx}\right)^{1/(b-a)}$. □

(The expression $(\mathscr{P}_a^b f(x)^{dx})^{1/(b-a)}$ does, in fact, represent a kind of "mean value." We will soon see that this expression is the natural way to define the geometric mean of f over the continuous interval $[a, b]$.)

We're finally ready to prove Part I of the Fundamental Theorem of Product Calculus.

Theorem 8. *(Fundamental Theorem of Product Calculus, Part I.) Let f be continuous and product-integrable on $[a, b]$, and let $F(x) = \mathscr{P}_a^x f(t)^{dt}$. Then F is continuous on $[a, b]$, product-differentiable on (a, b), and*

$$F^*(x) = f(x).$$

Proof. We have, for x in (a, b) and sufficiently small positive Δx,

$$\left(\frac{F(x+\Delta x)}{F(x)}\right)^{\frac{1}{\Delta x}} = \left(\frac{\mathscr{P}_a^{x+\Delta x} f(t)^{dt}}{\mathscr{P}_a^x f(t)^{dt}}\right)^{\frac{1}{\Delta x}} = \left(\mathscr{P}_x^{x+\Delta x} f(t)^{dt}\right)^{\frac{1}{\Delta x}}, \text{ by Lemma 5.}$$

By the Mean Value Theorem for Product Integrals, the value of the expression on the right is one of the values of f on $[x, x + \Delta x]$. Thus we have

$$\left(\mathscr{P}_x^{x+\Delta x} f(t)^{dt}\right)^{\frac{1}{\Delta x}} = f(c)$$

for some c in $[x, x+\Delta x]$. Since c must be between x and $x+\Delta x$, if we let Δx approach 0 then c must approach x. Since f is continuous, this means that $f(c)$ approaches $f(x)$ as Δx approaches 0 from the right. A similar argument shows that $f(c)$ approaches $f(x)$ for negative Δx, too; that is, if Δx approaches 0 from the left. Thus we have

$$F^*(x) = \lim_{\Delta x \to 0} \left(\frac{F(x+\Delta x)}{F(x)}\right)^{\frac{1}{\Delta x}} = \lim_{\Delta x \to 0} \left(\mathscr{P}_x^{x+\Delta x} f(t)^{dt}\right)^{\frac{1}{\Delta x}} = \lim_{\Delta x \to 0} f(c) = f(x).$$

To prove continuity of F, we know that F is positive on $[a, b]$. The existence of F^* on (a, b) then implies, by Theorem 4, that F is continuous on (a, b). If $x = a$ or b, then we can work through the argument we just made, where Δx is forced to be positive or negative, respectively, and the limit defining $F^*(x)$ is interpreted as the appropriate one-sided limit. The existence of this limit then shows, also via Theorem 4, that F is continuous at a and at b as well. □

Of course, it's the other part of the fundamental theorem that's useful for evaluating integrals. Here's the product calculus version.

Theorem 9. *(Fundamental Theorem of Product Calculus, Part II.) Let f be continuous and product-integrable on $[a, b]$, and let F be a function such that $F^*(x) = f(x)$ for all x in $[a, b]$. Then*

$$\mathscr{P}_a^b f(x)^{dx} = \frac{F(b)}{F(a)}.$$

Proof. Let $G(x) = \mathscr{P}_a^x f(t)^{dt}$. By Part I of the Fundamental Theorem of Product Calculus, $G^*(x) = f(x)$, and G is continuous on $[a, b]$. Since f is product-integrable, $f(x) > 0$ for all x in $[a, b]$, which, via Theorem 4, implies that F is continuous on $[a, b]$.

By Corollary 3, there exists some constant K such that $F(x) = KG(x)$ for all x in (a, b). As both F and G are continuous on $[a, b]$, taking appropriate one-sided limits then shows that $F(x) = KG(x)$ for $x = a$ and $x = b$ as well. Therefore,

$$\begin{aligned}
\frac{F(b)}{F(a)} &= \frac{KG(b)}{KG(a)} \\
&= \frac{\mathscr{P}_a^b f(x)^{dx}}{\mathscr{P}_a^a f(x)^{dx}} \\
&= \mathscr{P}_a^b f(x)^{dx}.
\end{aligned}$$

□

Product integrals and the usual integral

Thanks to the fundamental theorem, several basic product integration rules can be obtained simply by inverting the product differentiation rules discussed previously. We will not list these explicitly.

However, it is worth mentioning that just as the product derivative can be expressed in terms of the usual derivative, the product integral can be expressed in terms of the usual integral.

Theorem 10. *If $f(x) > 0$ for all x in $[a, b]$, then f is product-integrable on $[a, b]$ if and only if f is integrable on $[a, b]$, in which case*

$$\mathscr{P}_a^b f(x)^{dx} = \exp\left(\int_a^b \ln f(x) dx\right).$$

Proof. Suppose f is product-integrable on $[a, b]$. Then we have

$$\begin{aligned}
\ln \mathscr{P}_a^b f(x)^{dx} &= \ln \left(\lim_{||P|| \to 0} \prod_{k=1}^{n} f(c_k)^{\Delta x_k} \right) \\
&= \lim_{||P|| \to 0} \ln \prod_{k=1}^{n} f(c_k)^{\Delta x_k} \\
&= \lim_{||P|| \to 0} \sum_{k=1}^{n} \Delta x_k \ln f(c_k) \\
&= \int_a^b \ln f(x) dx.
\end{aligned}$$

Thus $\ln f(x)$ is integrable on $[a, b]$, and $\mathscr{P}_a^b f(x)^{dx} = \exp(\int_a^b \ln f(x) dx)$.

Since e^x is continuous, the integrability of $\ln f(x)$ on $[a,b]$ implies the integrability of $e^{\ln f(x)} = f(x)$ [1, p. 216].

If f is integrable on $[a,b]$, we can reverse these steps to show that f is product-integrable on $[a,b]$ and that $\mathscr{P}_a^b f(x)^{dx} = \exp(\int_a^b \ln f(x)dx)$. □

Applications of the product integral

What is the product integral useful for? As with the product derivative, we give a few applications. Several others are given in Dollard and Friedman [2].

Variable growth factors

A classic example of the use of exponential functions is in calculating the future value of an investment in which interest is compounded continually. If an amount A is compounded continually at an interest rate r, then the future value of the investment is given by $F(t) = Ae^{rt}$ [11, p. 44].

However, this formula assumes that the interest rate r is constant. Rarely do interest rates on investments remain constant. How would one determine the future value of an investment in which r is allowed to be a function of time? We can answer this question using the product integral.

If we want to calculate the future value at time b of an amount A invested at time a, we can slice up the interval $[a,b]$ into n subintervals small enough so that r is approximately constant over each subinterval. The future value at the end of the first subinterval is approximately $Ae^{r_1\Delta t_1}$, the future value at the end of the second subinterval is approximately $Ae^{r_1\Delta t_1}e^{r_2\Delta t_2}$, and after all n subintervals have passed the future value is approximately

$$F \approx A \prod_{k=1}^{n} e^{r(t_k)\Delta t_k}.$$

Using smaller subintervals will give a more accurate approximation, and letting the width of the largest subinterval approach zero should give the exact value. However, that's also the definition of the product integral. Thus we obtain

$$F = A \lim_{||P||\to 0} \prod_{k=1}^{n} \left(e^{r(t_k)}\right)^{\Delta t_k} = A\mathscr{P}_a^b \left(e^{r(t)}\right)^{dt}$$

for the future value of an investment in which the interest rate is $r(t)$ and interest is compounded continuously.

Since $e^{r(t)} > 0$, applying Theorem 10 yields a nice form for F in terms of the usual integral:

$$F = A \exp\left(\int_a^b r(t)dt\right). \qquad (1)$$

This approach works for any quantity in which the underlying process is multiplicative and has a variable growth or decay factor. While not as common as the corresponding additive processes on which the usual integral can be used they still occur regularly. For example, the product integral could be used to determine decay of charge on a capacitor in which capacitance or resistance is variable, to calculate population growth in which the growth factor is variable, or even in certain kinds of probability calculations – such as the probability of no arrivals in a Poisson process with a variable arrival rate. In addition, Guenther [5] discusses the use of the product integral to calculate the Helmholtz free energy of a photon gas, and Gill and Johansen [3] show how survival functions may be obtained by product-integrating hazard functions.

(Incidentally, (1) can also be obtained by solving the initial-value problem $\frac{dF}{dt} = r(t)F$, $F(a) = A$.)

Geometric mean

The *geometric mean* of a finite set $\{x_1, x_2, \ldots, x_n\}$ is given by

$$\sqrt[n]{x_1 x_2 \cdots x_n}.$$

What would it mean to take the geometric mean of a function over a continuous interval? For example, what is the geometric mean of the continuous set of numbers $[0, 1]$?

A reasonable approach to defining the geometric mean of f over $[a, b]$, paralleling the definition of the arithmetic mean of f over $[a, b]$, is the following: Partition $[a, b]$ into n equal subintervals. For each k, choose a point c_k to represent subinterval k. The expression $\sqrt[n]{\prod_{k=1}^{n} f(c_k)}$ then gives an approximation of what should be the geometric mean of f over $[a, b]$. But equal subintervals means $\Delta x_k = (b - a)/n$ for each k, which implies $1/n = \Delta x_k/(b - a)$. Thus our approximation of the geometric mean can be expressed as

$$\prod_{k=1}^{n} f(c_k)^{1/n} = \prod_{k=1}^{n} f(c_k)^{\Delta x_k/(b-a)} = \left(\prod_{k=1}^{n} f(c_k)^{\Delta x_k}\right)^{1/(b-a)}.$$

Taking the limit of this expression as the widths of the subintervals go to zero makes sense as the exact value of the geometric mean. But doing so also yields $(\mathscr{P}_a^b f(x)^{dx})^{1/(b-a)}$. A reasonable definition of the geometric mean, then, is the following:

Definition. *If f is product-integrable over $[a, b]$, then the geometric mean of f on $[a, b]$ is given by*

$$(\mathscr{P}_a^b f(x)^{dx})^{1/(b-a)}.$$

This definition is that obtained by changing usual integration to product integration and division to taking roots in the definition of the arithmetic mean, as the arithmetic mean of f over $[a,b]$ is given by

$$\frac{1}{b-a}\int_a^b f(x)\,dx.$$

We can now answer the question of the geometric mean of the set $[0,1]$, as this is the same as the geometric mean of the function $f(x)=x$ over $[0,1]$. However, we have to be a little careful because of the zero. In fact, $\mathscr{P}_0^1 x^{dx}$ is not defined under our definition of the product integral, precisely because of the zero. We can get around this by considering $\mathscr{P}_0^1 x^{dx}$ as an improper product integral and defining it to be $\lim_{a\to 0^+}\mathscr{P}_a^1 x^{dx}$, provided the limit exists. Since the product derivative of $(x/e)^x$ is x, we have

$$\lim_{a\to 0^+}(\mathscr{P}_a^1 x^{dx})^{1/(1-a)} = \lim_{a\to 0^+}\left(\frac{(1/e)^1}{(a/e)^a}\right)^{\frac{1}{1-a}} = \frac{1}{e}\approx 0.368,$$

as we know $\lim_{a\to 0^+} a^a = 1$. Thus the geometric mean of $[0,1]$ is $1/e$.

Incidentally, if we use $\mathscr{P}_1^n x^{dx}$ to approximate $n! = \prod_{k=1}^n k$, we obtain

$$n! \approx \frac{(n/e)^n}{(1/e)^1} = \frac{n^n}{e^{n-1}},$$

which, except for a factor of $\sqrt{n}$, is not far off from Stirling's well-known approximation for $n!$ [4, p. 452]:

$$n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$

Interestingly enough, the additive error when approximating $\sum_{k=0}^n k = \frac{n(n+1)}{2}$ by $\int_0^n x\,dx$ is $\frac{n}{2}$, the additive expression corresponding to $\sqrt{n}$.

Area on semilog graphs

Since the product derivative gives the slope of a curve plotted on a semilog graph, in a certain sense, it stands to reason that the product integral gives the area under a curve plotted on a semilog graph, again, in a certain sense. This is the case, but, once again, we need to make "in a certain sense" precise. To do so, we need the following definition of area corresponding to "geometric slope" and "geometric value." Given a region drawn on a graph, superimpose on the graph a linear scale identical on both vertical and horizontal axes. The area of the region measured by the linear scale is the *geometric area* of the region with respect to the linear scale. Like geometric value but unlike geometric slope, this quantity can change with the linear scale used, and thus its definition depends on the linear scale.

Then, *if $f(t)\geq 1$ for all t in $[a,x]$, the geometric value of $\mathscr{P}_a^x f(t)^{dt}$ gives the geometric area of the region under the graph of f, between $t=a$ and $t=x$, and* *above the horizontal axis, when f is plotted on a semilog graph, with respect to the linear scale defined by the horizontal axis of the semilog graph.* In other words, the area interpretation for the product integral is what we would expect based on the slope interpretation of the product derivative, including the fact that we need to retain the horizontal scaling from the original semilog graph.

We can see this from Theorem 10. Suppose $f(t) \geq 1$ for all t in $[a, x]$, and plot f on a semilog graph with b as the base of the logarithm. The geometric area under the curve of f from a to x and above the horizontal axis, with respect to the horizontal scaling, is given by the area function

$$A(x) = \int_a^x \log_b f(t)dt = \frac{1}{\ln b} \int_a^x \ln f(t)dt.$$

If we plot $b^{A(x)}$ on the same semilog graph we will obtain a curve whose geometric value gives the geometric area under the graph of f. But

$$b^{A(x)} = b^{\frac{1}{\ln b} \int_a^x \ln f(t)dt} = \left(e^{\ln b}\right)^{\frac{1}{\ln b} \int_a^x \ln f(t)dt} = e^{\int_a^x \ln f(t)dt} = \mathscr{P}_a^x f(t)^{dt}.$$

If $0 < f(x) < 1$, then the plot of f on a semilog graph will dip below the horizontal axis, and $\mathscr{P}_a^x f(t)dt$ can then be interpreted as a geometric area above the horizontal axis less a geometric area below the horizontal axis (just as with the usual integral).

For example, Figure 8 contains a plot of $f(x) = x$ and its area function $\mathscr{P}_1^x f(t)dt = e(x/e)^x$ on a semilog graph. (The extra factor of e comes from the constant of product integration, as two functions with the same product derivative are equal up to a constant factor.) The curve of $e(x/e)^x$ (the concave up graph) behaves as the geometric area

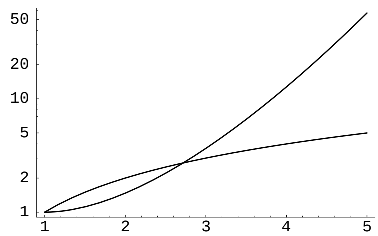

Figure 8: Semilog plots of x and $e(x/e)^x$

function for the second curve, with respect to the horizontal scaling on the graph.

As with the derivatives, the geometric values of the usual integral of f and the product integral of f can both be thought of as yielding the geometric area under the graph of f; the difference is that the scaling on the original graph of f is linear-linear for the usual integral and is semilog for the product integral.

Conclusions

My hope in presenting this material is that doing so makes product calculus more accessible to undergraduates. Nearly every result or application mimics a similar result or application with the usual calculus. Because of this, comparing and contrasting the results here with those of the usual calculus helps bring out the essential aspects of the calculus – in other words, what makes calculus calculus. Interested students may be able to make projects out of taking aspects of the usual calculus not touched on here and working out their product calculus versions as well.

Simply knowing that product calculus exists may also be useful in helping students understand how crucial definitions are in mathematics. I find it quite difficult to communicate this importance to freshmen. Saying it doesn't help much without examples; perhaps product calculus can serve as an example. This seemed to be the case with Sarah, in fact. After the conversation recounted at the beginning of the article she never asked me again about the meaning of the derivative. She seemed to have grasped at that point that its meaning is entailed in whatever the definition implies. After all, changing the definition, even slightly, can yield a different kind of calculus, so it must be the definition that really matters.

This material raises some research questions, too. Virtually all of the product calculus results can be obtained simply by making appropriate changes to results and proofs for the usual calculus. This implies that what makes the usual calculus work are certain relationships between the operations in the definitions of the derivative and the integral. We have shown that making a set of changes to these definitions that preserves these relationships produces another kind of calculus. Might other changes produce other kinds of calculus? For example, what changes could be used to produce derivatives that yield Taylor-like theorems for other kinds of functions? What changes could produce integrals that could be used to compute other kinds of means? Maybe there are an infinite number of kinds of calculus, one for each way of defining the mean. And if there are more kinds of calculus, is there is a general theory of calculus that covers all of them?

Acknowledgments Thanks to Martin Jackson, Alison Paradise, and David Scott for comments on the original draft of the paper. Extra special thanks to Sarah Glancy for tenacity and for asking good questions.

References

[1] Robert G. Bartle and Donald L. Sherbert, *Introduction to Real Analysis*, 3rd ed., John Wiley & Sons, 2000.

[2] John D. Dollard and Charles N. Friedman, *Product Integration, with Applications to Differential Equations*, Addison-Wesley, 1979.

[3] Richard D. Gill and Søren Johansen, A survey of product-integration with a view toward application in survival analysis, *Annals of Statistics* **18** (1990), 1501–1555.

[4] Ronald L. Graham, Donald E. Knuth, and Oren Patashnik, *Concrete Mathematics: A Foundation for Computer Science*, Addison-Wesley, 2nd ed., 1994.

[5] Raymond A. Guenther, Product integrals and sum integrals, *International Journal of Mathematical Education in Science and Technology* **14** (1983), 243–249.

[6] David Manura, An introduction to generalized calculus, `http://www.math2.org/math/paper/contents.htm`.

[7] Math message boards, `http://www.math2.org/mmb/`.

[8] Monty J. Strauss, Gerald L. Bradley, and Karl L. Smith, *Calculus*, 3rd ed., Prentice Hall, 2002.

[9] Vito Volterra, Sulle equazioni differenziali lineari, *Rendiconti Accademia dei Lincei* (Series 4) **3** (1887), 393–396.

[10] B. Wang, C. Molina, and E. Abdalla, Evolution of a massless scalar field in Reissner-Nordström anti-de Sitter spacetimes, *Physical Review D* **63** (2001), Article 084001.

[11] Maurice D. Weir, Joel Hass, and Frank R. Giordano, *Thomas' Calculus: Early Transcendentals*, Pearson Addison Wesley, 2006.