engineers to see the light and consider mathematics as a career, though only time will tell.

Clearly, this experimental approach to investigating advanced concepts and gaining insights into basic concepts has much to recommend it. I can see the same idea being applied in several other freshman and sophomore level courses. For example, in a linear algebra course it could lead students into an experimental investigation of the concepts of functional analysis, or in a multivariable calculus course it could lead to a computer-aided differential geometry project. Such projects could be implemented using any of the existing popular mathematical software packages, or even developed from scratch with computer language compilers. Even the development of such a project could become a project for an advanced student with sufficient computer expertise. I hope others will try these ideas; and if they do, I trust they will experience as much success as I did.

A Natural Proof of the Chain Rule

Stephen Kenton (kenton@ecsuc.ctstateu.edu), Eastern Conn. State Univ., Willimantic, CT 06226

The first three editions of Hardy's *A Course in Pure Mathematics* **[4]** contain a "natural" proof of the familiar Chain Rule for differentiating the composition of two real-valued functions of a real variable. Unfortunately, the proof was wrong!

Chain Rule. Let I, J be open intervals of real numbers, $f: I \to J, g: J \to \Re$, f differentiable at c, g differentiable at $f(c)$. Then $g \circ f$ is differentiable at c, and $(g \circ f)'(c) = g'(f(c)) \cdot f'(c)$.

The undergraduate real analysis (or advanced calculus) course allows students to experience the striking power of creating and proving significant results by making natural choices and educated guesses, and then proving these results using a few basic techniques. For the Chain Rule proof, one begins with the definition of a derivative, a familiar technique is applied to transform the problem so the hypotheses can be used, and then the proof follows easily. But then a subtle flaw is revealed, the attempt abandoned, and a special technique is introduced; e.g., **[2–7]**. Often, the motivational step is skipped and the unmotivated proof is directly presented; e.g., **[1]**, **[8]**. In this note, we show how to prove the Chain Rule by following the original path, using techniques familiar to students from previous work.

Natural "proof" of the Chain Rule. Using the definition of the derivative of a composite function,

$$(g \circ f)'(c) := \lim_{x \to c} \frac{g(f(x)) - g(f(c))}{x - c} \tag{1}$$

we rewrite the fraction in terms of difference quotients for the two functions being composed:

$$\lim_{x \to c} \frac{g(f(x)) - g(f(c))}{f(x) - f(c)} \frac{f(x) - f(c)}{x - c}. \tag{2}$$

Recognizing that f is continuous at c (so $f(x) \to f(c)$ as $x \to c$), the result follows. However, this argument has a flaw which Hardy points out in later editions of *A Course of Pure Mathematics* [4], "The proofs in many text-books (and in the first three editions of this book) are inaccurate... The error was in overlooking the possibility that $f(x) = f(c)$."

Indeed, the argument given above proves the Chain Rule in the case in which there is a neighborhood of c inside which the difference $f(x) - f(c)$ has no zeros other than c. But, suppose the worst: *$f(x) - f(c)$ has zeros distinct from c in every neighborhood of c.* This immediately allows us to construct an infinite sequence of points, $x_1, x_2, \ldots, x_n, \ldots$ such that for each n, $f(x_n) = f(c)$ and the nonzero distance $|x_n - c| < \frac{1}{n}$. (As an illustration, one might think of the fundamental counter-example

$$f(x) = \begin{cases} x^2 \sin \dfrac{1}{x}, & x \neq 0 \\ 0, & x = 0 \end{cases}$$

when $c = 0$.) Thinking of this example leads to the following surprising result: Under the additional condition that $f(x) - f(c)$ has zeroes distinct from c in every neighborhood of c,

$$f'(c) := \lim_{x \to c} \frac{f(x) - f(c)}{x - c} = \lim_{x_n \to c} \frac{f(x_n) - f(c)}{x_n - c} = 0. \tag{3}$$

The proof will be complete if we can also show $(g \circ f)'(c) = 0$. Note that we cannot use the same type of proof as in (3) for this assertion since we do not know that the limit in (1) exists.

To prove $(g \circ f)'(c) = 0$, we show the *supremum* of the fraction in (1) can be made arbitrarily small in sufficiently small neighborhoods of c. For this purpose, we can ignore all the points x such that $f(x) = f(c)$, thus avoiding the problem upon which the original proof floundered!

$$\begin{aligned} &\sup\left(\left|\frac{g(f(x)) - g(f(c))}{x - c}\right| : 0 < |x - c| < \delta\right) \\ &= \sup\left(\left|\frac{g(f(x)) - g(f(c))}{x - c}\right| : 0 < |x - c| < \delta, f(x) \neq f(c)\right) \\ &= \sup\left(\left|\frac{g(f(x)) - g(f(c))}{f(x) - f(c)} \frac{f(x) - f(c)}{x - c}\right| : 0 < |x - c| < \delta, f(x) \neq f(c)\right). \end{aligned}$$

which we recognize as the difference quotient in (2). The first fraction in the last expression is bounded in some neighborhood of c since g is differentiable at $f(c)$, and the second fraction can be made arbitrarily small since $f'(c) = 0$. This shows that $\frac{g(f(x))-g(f(c))}{x-c}$ can be made arbitrarily small in a sufficiently small neighborhood of c, so the defining limit exists and is zero, proving the theorem.

References

1. T. Apostol, *Mathematical Analysis* (2*nd ed.*), Addison Wesley, 1975.
2. R. Bartle and D. Sherbert, *Introduction to Real Analysis* (2*nd ed.*), Wiley, 1992.
3. G. Bilodeau and P. Thie, *An Introduction to Analysis*, McGraw-Hill, 1997.
4. G. Hardy, *A Course in Pure Mathematics* (10*th ed.*), Cambridge University Press, 1950.
5. W. Kosmala, *Introductory Mathematical Analysis*, Wm. C. Brown, 1995.
6. S. Lay, *Analysis with an Introduction to Proof* (2*nd ed.*), Prentice Hall, 1990.
7. J. Olmsted, *Intermediate Analysis*, Appleton-Century-Crofts, 1956.
8. W. Rudin, *Principles of Mathematical Analysis* (3*rd ed.*), McGraw Hill, 1976.

---o---

The Average Distance of the Earth from the Sun

David L. Deever (ddeever@otterbein.edu) Otterbein College, Westerville OH 43081

In astronomy texts one sometimes finds a description of an Astronomical Unit (AU) as the "average distance of the earth from the sun." However, consideration of the concept of the average, or mean value, of a function reveals that the parameterization (or, equivalently, the underlying metric) must be specified in order to have a precise definition of that mean value. A more careful definition of AU is as the length of the semi-major axis of the earth's orbit. In this note we will show how the choice of parameter from among several plausible alternatives can affect the average value of the distance.

In order to investigate this topic we recall that the average value of the function f over the interval $[a, b]$ is given by

$$Avg(f, t) = \frac{1}{b-a}\int_a^b f(t)\, dt \tag{1}$$

and that, if $\tau = \varphi(t)$ is a change of parameter with $\alpha = \varphi(a)$ and $\beta = \varphi(b)$, then

$$Avg(f, \tau) = \frac{1}{\beta-\alpha}\int_{\tau=\alpha}^{\tau=\beta} f(\varphi^{-1}(\tau))\, d\tau = \frac{\int_{t=a}^{t=b} f(t)\,\varphi'(t)\, dt}{\int_{t=a}^{t=b} \varphi'(t)\, dt} = \frac{\int_{t=a}^{t=b} f(t)\, d\tau}{\int_{t=a}^{t=b} d\tau}$$

gives its value with respect to the new parameter.

In the following examples we will consider the polar equation of an ellipse,

$$r = \frac{1-e^2}{1+e\cos\theta} \tag{2}$$