Things I Learned Teaching AP Calculus

Chuck Garner, Ph.D.
Rockdale Magnet School for Science and Technology

October 17, 2014

I started teaching AP Calculus in 2002 at a magnet school for science and technology. I was not very good at it, which was a shame, since every student in my school had to take calculus to graduate, and I was the only calculus teacher. I realize now why I was not good at it. I taught calculus to my students the way I was taught calculus: heavy on symbolic manipulation and algebra skills; more concerned with finding antiderivatives than with understanding what the definite integral means; and only focused on those applications which could be approached in a rote fashion.

I have learned a lot since those first years teaching calculus. By focusing more on the conceptual rather than the computational, my students understand calculus better and have greater success on the AP exam. This approach helped my students, but it also helped me. I have stumbled across some interesting insights over the years, some developed over time and others gained in flashes of insight. You may already be familiar with much of what I have learned. But if you still teach calculus the way you were taught, you may learn something too. So indulge me as I present to you some of the things I have learned teaching AP Calculus.

The first thing I learned was that we should

Keep the algebra simple.

When I first started teaching AP Calculus, I went absolutely nuts trying to get kids to understand how to evaluate expressions like

$$\lim_{x\to 1}\frac{3x^3+7x^2-2x-8}{(x-1)^5},\ \lim_{x\to 3}\frac{\sqrt{x+3}-\sqrt{6}}{x-3},\ \text{and}\ \int\frac{x^4-2}{x^4-2x^3+5x^2-8x+4}\,dx.$$

So much effort and frustration and time spent on the algebra! Even though most students recognized what to do, most students struggled to perform the task. But in studying the types of problems which actually appear on the AP Calculus

Exam, I realized that no problems like these are tested. My guess as to why problems such as these do not appear is that students should not be inhibited from demonstrating knowledge of calculus because of the algebra.

But more than that, I would ask whether they are instructive at all. How can those three problems above be more instructive to students than

$$\lim_{x\to 2}\frac{x^2-4}{x-2} \qquad \text{or} \qquad \int \frac{4x^2}{x^2-4}\,dx,$$

two problems that do not require great amounts of tedious algebra. Besides, is this tedious algebra really necessary in our modern world? Particularly when many practicing scientists and engineers would use a computer to calculate anything which may involve such awful algebraic manipulations.

I can't believe how I made my students work at the beginning of each year. I covered the typical course sequence: those first two weeks of algebra review, followed by weeks of limits, before getting into the good stuff. And we require them to do crazy algebra to evaluate limits in order to illustrate *why* they absolutely had to have those algebra skills down! (As if we are saying "See! Limits are an application of all those algebra skills! You must know algebra!") I spent so much time working with limits and getting kids to evaluate them using algebra tricks. However, once I examined the types of limit problems that are asked on the AP exam, I realized that

Limits are a waste of time.

I know this is a controversial statement, but it is true: limits are a waste of time. Now, I'm not advocating the overthrow of established calculus hegemony. I realize the need for the epsilon-delta formulation as much as the next mathematician. However, the epsilon-delta form of limits has been long absent from the AP Calculus course, and the limit problems which remain all seem to fall into four categories: (1) direct evaluation; (2) left- and right-hand limits for piecewise functions; (3) definitions of continuity and the derivative; and all other limits can be solved with (4) l'Hôpital's rule.

The order of these four categories is the order in which I present these concepts about limits. I teach just enough about limits—through direct evaluation, piecewise functions, and functions whose graphs have holes—for students to understand the definition of the derivative. The key explanation which allows such an abbreviated treatment of limits is that I relate "limit" to the simple idea of "behavior." Thus, a limit describes a function's behavior near a point.

For example: What is the behavior of $f(x) = (x^2 + 2x)/x$ near $x = 0$? Since the function $g(x) = x + 2$ is exactly equal to $f(x)$ for all $x \neq 2$, we may use g to model the behavior of f. Since the behavior of g near $x = 0$ is 2, the behavior of f near $x = 0$ must also be 2. That is, f is behaving like it will be 2 as x gets closer to zero. Thus, the limit of $f(x)$ as $x \to 0$ is 2.

After this, I continue with the standard series of lessons on differentiation, and conclude that with l'Hôpital's rule so students can handle all other kinds of limits, including limits involving infinity. Removing the need for all the ridiculous algebra has made my life and my students' lives much better, without sacrificing understanding. The massive amount of algebraic simplification has been reduced. This, along with the AP exam standard of not needing to simplify arithmetic, has lead me to a simple rule:

Don't simplify.

As AP Calculus teachers, we are already familiar with the fact that fractions need not be reduced. And even some arithmetic need not be performed. Certainly, however, some simplification is just never needed. I mean, why should anyone anywhere rationalize any denominator? (Can't math teachers agree that $\sin(\pi/4) = 1/\sqrt{2}$ is a perfectly good answer already?) Rationalization is nothing more than an antiquated holdover from the days before calculators. Let me explain: Notice that it is impossible to use long division to divide 1 by $\sqrt{2} = 1.41421356237\ldots$. But when you rationalize, then you can divide $\sqrt{2}$ by 2, to whatever degree of accuracy you need. But now we have calculators, and we no longer need to rationalize.

Even some algebraic simplification is unwarranted. Consider finding the domain of a function. If one simplifies first, then you may miss some values on the domain. This function appeared on the AP exam about 15 years ago:

$$f(x) = \sqrt{x^4 - 16x^2}$$

The first question asked the student to find the domain. If the student simplified it to $x\sqrt{x^2 - 16}$, then the domain is changed. (Not to mention the proper expression is $|x|\sqrt{x^2 - 16}$, but we can't expect all our students to be miracle workers!) The domain of the original function includes zero, but the domain of the simplified form excludes zero. This is a mistake, as the simplified form actually leads to a wrong answer! The correct domain is, of course, $(-\infty, -4] \cup \{0\} \cup [4, \infty)$.

Don't think simplifying expressions is never needed. Simplification is needed for those multiple-choice problems—students must match what they have obtained to the correct answer choice. In addition to algebra, students must be able to simplify expressions involving trigonometric functions, logarithmic functions, and exponential functions. (Which, by the way, is not strictly considered simplification. This would fall under *evaluating* trigonometric, logarithmic, and exponential functions, which must always be done, when it is possible to do so.) The logarithmic functions and their associated properties are particularly important. Unfortunately, many students, if they remember those "log properties" at all, do not know why they work or even why they exist. But there is a way to show them why, and that's by waiting to introduce them in calculus until integration. This way, you can

Define the logarithm as an integral.

What I present to my students—after teaching them the Fundamental Theorem—is a variation of the following. I define the function $L(x)$ as

$$L(x) = \int_1^x \frac{1}{t}\,dt.$$

I do not tell them that this is the logarithm; instead we (the class and I) will derive all the properties of the logarithm from this definition. This way, the properties have a reason for existing.

Notice that $L(x)$ cannot be defined for zero or negative values, because of the vertical asymptote of the function $f(t) = 1/t$. This function L is defined in this way because its derivative needs to be $1/x$, which it is, according to the Fundamental Theorem:

$$L'(x) = \frac{d}{dx}\int_1^x \frac{1}{t}\,dt = \frac{1}{x}.$$

Hence, $L(x)$ is the antiderivative of $\frac{1}{x}$. Now we investigate properties of L. First, the class sees that

$$L(1) = \int_1^1 \frac{1}{t}\,dt = 0.$$

Clearly, if $x > 1$ then $L(x)$ is positive since we are calculating area under a curve. Also, if $0 < x < 1$, then, since the integral accumulates area in the opposite direction, $L(x)$ must be negative.

Is L the only function for which $1/x$ is its derivative? We answer this by finding the derivative of $L(kx)$ for constant k. By the chain rule,

$$L'(kx) = \frac{d}{dx}\int_1^{kx} \frac{1}{t}\,dt = \frac{1}{kx}\cdot k = \frac{1}{x}$$

so that $L(kx)$ is also an antiderivative of $1/x$. Hence, since two antiderivatives can at most differ by a constant, $L(kx) = L(x) + C$. We may find the value of C by letting $x = 1$, where this becomes $L(k) = L(1) + C$. But $L(1) = 0$, which implies $L(k) = C$. Therefore,

$$L(kx) = L(x) + L(k).$$

I believe this is the development of the "log property" $\log(ab) = \log a + \log b$ that makes the most sense to students at this level.

Then the class differentiates $L(x^p)$ for real p:

$$L'(x^p) = \frac{1}{x^p}\cdot px^{p-1} = p\cdot\frac{1}{x} = pL'(x).$$

So then $p(1/x)$ is the antiderivative of the two functions $L(x^p)$ and $p\cdot L(x)$. Hence, these functions must differ by a constant; this gives $L(x^p) = pL(x) + C$. Letting $x = 1$ results in $C = 0$. Therefore, in general,

$$L(x^p) = pL(x).$$

At this point, I finally define the function $L(x)$ as the logarithm function for my students, and point out that there must be some value of x which results in $L(x) = 1$. I simply say that whatever value that is we will call e.

Then we investigate functions of the form $f(x) = cL(x)$ for some nonzero constant c. All properties so far discovered must be satisfied; in particular, we want $cL(x) = 1$ for a particular value of x. Call this particular value b. Then b cannot be equal to 1 (since $L(1) = 0$), and so we have $cL(b) = 1$. Rewriting, we get $c = 1/L(b)$, so that $f(x) = L(x)/L(b)$. This implies that all functions which are multiples of $L(x)$ are simply $L(x)$ scaled by a particular value of L. Note that when this value is e, we have

$$\frac{L(x)}{L(b)} = \frac{L(x)}{L(e)} = \frac{L(x)}{1} = L(x),$$

which is the reason we term this "non-scaled" function the *natural logarithm*. Writing the natural logarithm of x as $\log_e x$ (to denote that the constant scale factor is $L(e)$), motivates us to define the other scaled functions of $L(x)$ as *logarithms to the base b*, denoted $\log_b x$, where b is any number greater than 1. Hence, with this new notation,

$$\frac{L(x)}{L(b)} = \log_b x.$$

However, $L(x)$ alone can be written $\log_e x$ and similarly $L(b)$ may be written as $\log_e b$. Thus,

$$\log_b x = \frac{L(x)}{L(b)} = \frac{\log_e x}{\log_e b},$$

and the change-of-base formula is established. And, of course $\log_e x$ is denoted $\ln x$.

With this established it is then an easy matter to discuss the properties of the inverse of the function $L(x)$, which I oh-so-cleverly call $E(x)$. Naturally, I lead the students to the fact that the inverse is the exponential function. But before we get there, I exploit the fact that $E(x)$ is the inverse of $L(x)$ in order to find the derivative of $E(x)$.

I begin by composing $\ln(x)$ with $E(x)$ in two ways. The first way is $\ln(E(x)) = x$. Recalling that $\ln(x) = \int_1^x \frac{1}{t}\,dt$, we also have that

$$\ln(E(x)) = \int_1^{E(x)} \frac{1}{t}\,dt.$$

Therefore,

$$\int_1^{E(x)} \frac{1}{t}\,dt = x.$$

Taking derivatives of both sides, we get

$$\frac{1}{E(x)} \cdot E'(x) = 1, \qquad \text{or} \qquad E'(x) = E(x).$$

This is the simplest way of actually proving that the derivative of the exponential function is itself. Only the development of the logarithm as an integral allows this. When we introduce exponentials and logarithms early in the AP Calculus course, we lose this mind-blowing result—instead, it becomes something only accessible through calculator tricks and numerical data. Which is a fine method if no other method is available; but there is, I think, a better method in this case.

Of course when it comes to exponentials, the traditional method of writing the inverse of the logarithm is by denoting it e^x. This is a fine notation—as far as it goes. However, when introduced using this notation, it becomes too easily confused in the minds of students with a power function. I know we have all seen the following ridiculous work in our students:

$$\frac{d}{dx}\left(e^{x^2}\right) = x^2 e^{x^2-1}. \qquad \text{(This is wrong!)}$$

How can we get our students to stop making this silly mistake? By never using this notation until later in the course. Instead

Use the notation exp(x) for the exponential function.

When students are asked to differentiate $f(x) = \cos\left(x^2\right)$, $g(x) = \ln\left(x^2\right)$ and $h(x) = \sin\left(x^2\right)$, the students do not use the power rule by mistake! They of course use the chain rule: $f'(x) = -2x\sin\left(x^2\right)$, $g'(x) = 2x(1/x^2) = 2/x$, and $h'(x) = 2x\cos\left(x^2\right)$. So why not use a notation for the exponential function which reinforces the fact that it is a function like these others? With this notation, a student will correctly obtain

$$\frac{d}{dx}\left(\exp\left(x^2\right)\right) = 2x\exp\left(x^2\right).$$

Like the logarithm function, the trigonometric functions, and the inverse trigonometric functions, we should use a notation that screams "function!"

What about the other exponential functions like 2^x and 3^x and the like? Well, since each of these can be expressed in terms of the function $\exp(x)$, like so—

$$2^x = \left(e^{\ln 2}\right)^x = e^{x\ln 2} = \exp\left(x\ln 2\right)$$

—one can avoid using any other exponential function but $\exp(x)$, just as one can avoid using any other logarithm function but $\ln x$.

Speaking of avoiding troublesome notations and functions, I have a bone to pick with the notation for inverse trigonometric functions. I can't stand it. That's why I

Use the "arc" notation for inverse trig functions.

I cannot count how many times students see $\sin^{-1} x$ and think "$\csc x$." It is ingrained in their little heads that an exponent of negative one must mean reciprocal. I have tried to unteach that fact, but unteaching is really hard to do! So I have resorted to using "arcsin x" for the inverse sine to avoid any confusion with the cosecant.

I much prefer this notation for another reason as well: this notation indicates what this function does. We know that we need an angle input for the sine function; that is obvious. But what do get out? When we write $\sin(\pi/6) = 1/2$, what exactly does that $1/2$ represent? It represents a length. Indeed, given a circle of radius 1, $\sin\theta = k$ says that the length of the chord determined by the angle 2θ is $2k$. In ancient times, the value of the length of the chord was used to denote the length of the arc subtended by the chord. That is, by constructing a chord of length $2k$, one then determines an arc, so the length of the chord became associated with the length of the arc. Therefore, the sine function takes an angle and from that one may determine the length of an arc. It makes sense that the inverse sine function should take a length of an arc and return the associated angle; hence "arcsine." I tell my students simply that sin(angle) = length and arcsin(length) = angle.

So not only does this notation resolve the confusion between $\sin^{-1} x$ and $\csc x$, this resolves most of the domain and range issues students have with these functions as well.

Those domain and range issues are tricky, but getting a handle on them is important! Most students ignore domain issues, particularly when a function is defined on an interval. This is especially important when finding extrema, which is why I

Use the "closed interval extrema test."

I teach the two classic tests for extrema: the first derivative test and the second derivative test. (The second derivative test is a test for extrema, not inflection points—there is an "inflection point test" for that.) Most AP Calculus teachers are good with teaching their students these tests. Most are also good with simply evaluating the function at the critical points and at the endpoints when the function is defined on a closed interval. But the students do not so a good job of catching the closed interval problems. By teaching this as a separate test, students are keyed into noticing whether the function is defined for an interval or not.

I teach many things not mentioned in the AP Calculus AB course description, such as l'Hôpital's rule. I am careful about the "extra" stuff I teach. I feel anything extra must make sense to the students and there must be a reason for its inclusion. I also teach things not mentioned in the AP Calculus BC course description. In fact, I

Teach the root test for convergence.

The root test for infinite series says the following.

Theorem 1 (Root Test). *For the series $\sum_{n=1}^{\infty} a_n$, define*

$$r = \lim_{n\to\infty} \sqrt[n]{|a_n|}.$$

If $r < 1$, then $\sum a_n$ converges absolutely; if $r > 1$, then $\sum a_n$ diverges.

Any series which can be shown to converge by the ratio test also can be shown to converge by the root test. (In fact, both the ratio and root tests have the same conclusion.) As long as one establishes that $\lim_{n\to\infty} \sqrt[n]{n!} = \infty$,[1] then the root test is, in my opinion, easier to use than the ratio test. For example, most of us would use the ratio test to solve the following problem.

Problem 1. *Determine the convergence or divergence of the series*

$$\sum_{n=0}^{\infty} \frac{(n+1)^n}{n!n^n}.$$

The ratio test certainly works, but the root test determines the convergence more quickly:

$$\lim_{n\to\infty} \sqrt[n]{\left|\frac{(n+1)^n}{n!n^n}\right|} = \lim_{n\to\infty} \frac{n+1}{n\sqrt[n]{n!}} = 0$$

so that the series converges absolutely. The root test is easy to apply since it requires less algebraic manipulation than the ratio test. (I still teach the ratio test since there have been AP Exam problems which specifically ask the students to use the ratio test.) Reducing algebraic manipulation is important for most of our students, since the algebra skills they do have are generally very weak. That is the motivation for why I

Use partial derivatives for implicit differentiation.

When confronted with an implicitly-defined curve, we AP Calculus teachers generally teach students to interpret each y as nothing more than representing an implicit function $y = f(x)$ so that the chain rule must be used on each y term. For instance, to find dy/dx for the curve $y^2 + x^3y^3 + x^2 - 8 = 0$ we find $2yy' + 3x^2y^3 + 3x^3y^2y' + 2x = 0$. Solving for y' we obtain

$$y' = \frac{-3x^2y^3 - 2x}{3x^3y^2 + 2y}.$$

The difficulty is that students forget to use the chain rule when differentiating y. This becomes more burdensome if the equation of the curve has many terms involving products or quotients of x and y. However, if we view the curve as

[1] This can be seen to diverge by splitting the expression up: $\sqrt[n]{1} \cdot \sqrt[n]{2} \cdot \sqrt[n]{3} \cdots \sqrt[n]{n}$. Each term converges to 1 from above, so this product consists of numbers each larger than 1, except for $\sqrt[n]{1} = 1$. Thus the product continues to grow as $n \to \infty$.

defining a surface in space, then we may use partial derivatives to easily obtain dy/dx.

Suppose $F(x, y) = 0$ where F is a function of the two independent variables x and y. Then the total differential of this function is

$$\frac{\partial F}{\partial x}dx + \frac{\partial F}{\partial y}dy = 0.$$

(See [2].) Solving this differential for the ratio dy/dx gives

$$\frac{dy}{dx} = -\frac{\partial F/\partial x}{\partial F/\partial y}.$$

Let's use this technique on the implicit function $F(x, y) = y^2 + x^3y^3 + x^2 - 8$. We have

$$\frac{dy}{dx} = -\frac{\partial F/\partial x}{\partial F/\partial y} = -\frac{3x^2y^3 + 2x}{2y + 3x^3y^2},$$

which is the same derivative we found before, but with much less trouble (and no chain rule errors)!

I think it is interesting that a change in notation, or a change in method, or a change in definitions, can lead to deeper understanding and easier execution. I personally had such an epiphany about eight years ago when I realized that

Euler's method is just tangent lines.

While my class was reviewing for the AP Exam, I happened to have done on the board a problem about tangent lines and another problem on Euler's method. I saw my Euler's method equation

$$y_{k+1} = \left.\frac{dy}{dx}\right|_{(x_k,y_k)} \Delta x + y_k$$

and I rewrote it:

$$\left.\frac{dy}{dx}\right|_{(x_k,y_k)} \Delta x = y_{k+1} - y_k.$$

I teach my students to write tangent lines in point-slope form:

$$m(x - x_1) = y - y_1$$

and I noticed that Euler's method is just tangent lines. The slope at a point is clearly $m = dy/dx|_{(x_k,y_k)}$ and the change in x, Δx, is $x - x_1$. This completely changed how I teach Euler's method. I teach this as nothing more that an application of tangent lines instead of as a strict numerical procedure. This approach eliminates one more equation for a student to memorize.

Note that Euler's method is itself an approximation for the value of a function. This is also the use of a tangent line: it too is an approximation. Tangent lines are very useful approximators. In fact,

Taylor polynomials are just an extension of the tangent line idea.

I introduce the concept of "approximating polynomials" early in the course. Once tangent lines are used to approximate functions, it is natural to ask how to get a better approximation. One answer is to add a quadratic term. This would match the concavity of the function. So we should use something like the following, to approximate $f(x)$ centered at $x = a$:

$$P_2(x) = A + B(x-a) + C(x-a)^2.$$

How do we know what the coefficients A, B, and C are? Well, certainly the values of f and P_1 must be the same at $x = a$. Therefore, we need $f(a) = P_2(a) = A$. Also, we want P_2 and f to be tangent; thus, we need $f'(a) = P_2'(a)$. Computing $P_2'(x)$ we get $P_2'(x) = B + 2C(x-a)$ so that $P_2'(a) = B = f'(a)$. This means that so far we have

$$P_2(x) = f(a) + f'(a)(x-a) + C(x-a)^2$$

and all that remains is to determine C.

We added the quadratic term so the approximating polynomial matches the concavity; so we need the *concavity* of $P_2(x)$ to match that of $f(x)$ at $x = a$. This implies that we need $f''(a) = P_2''(a)$. Thus $P_2''(x) = 2C$ so then $P_2''(a) = 2C = f''(a)$; this implies $C = f''(a)/2$.

At this point, I define the "quadratic approximating polynomial" for f at $x = a$ to be

$$P_2(x) = f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2$$

where a is the center.

And of course, it is natural to ask about a cubic approximation. In fact, we can determine such an approximation using previous methods. Here, we want to determine coefficients A, B, C, and D such that

$$P_3(x) = A + B(x-a) + C(x-a)^2 + D(x-a)^3$$

where, again, $x = a$ is the center. We will end up going through the same procedure as with the quadratic approximation to obtain A, B, and C. Thus, we know $A = f(a)$, $B = f'(a)$, and $C = f''(a)/2$. It only remains to determine D. Following the pattern established, it makes sense that we need the third derivatives of f and P_3 to be equal at $x = a$. Computing, we have

$$\begin{aligned} P_3'(x) &= B + 2C(x-a) + 3D(x-a)^2 \\ P_3''(x) &= 2C + 6D(x-a) \\ P_3'''(x) &= 6D \end{aligned}$$

so then $P_3'''(a) = f'''(a) = 6D$, or $D = f'''(a)/6$. And, we have a definition of a "cubic approximating polynomial" of a function f at $x = a$ as

$$P_3(x) = f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + \frac{1}{6}f'''(a)(x-a)^3$$

where a is the center.

Discussions of convergence or divergence are held until Taylor series and Taylor polynomials are introduced later. However, this approach sets up the idea of Taylor series and Taylor polynomials. The students already understand the use of Taylor polynomials as approximators and can use them as such. The general form of a Taylor polynomial also makes sense, as students have seen and understood why it has that form. Many modern calculus books do not use this approach, although some old ones do. And a few old books use this idea very early in the course. This leads to one thing I would encourage every calculus teacher to do:

Read old calculus textbooks.

Older calculus textbooks contain gems not contained in new textbooks. (Why is that? Do mathematical terms fall out of favor? Are certain symbols and notations just a passing fad? Are modern textbooks too concerned with providing support for technology to actually explain things well? I wish I knew.) The notation "$\exp(x)$" for the exponential function I found in an old calculus book (see [1] for example). Older calculus books were very concerned with curve-sketching—makes sense that they would be in the era before graphing calculators. However, notice that today in our calculus classes we do not spend any time at all sketching implicitly defined curves, even though our graphing calculators cannot graph them. So it would seem to me that knowing how to sketch implicitly-defined curves would continue to be useful to our students. Older calculus textbooks are filled with such ideas, while newer ones are not. I agree that technology should be used to graph these kinds of curves, but by disregarding all aspects of implicit curve-sketching, we lose some pertinent ideas for AP calculus. Allow me to describe one example of such a thing. Consider the following.

Problem 2. *Find all points on the curve $y^4 - 5y^2 = x^4 - 4x^2$ at which the tangent lines are (a) vertical; (b) horizontal.*

Most of us would (quite rightly) find the derivative first. Let's use partial derivatives to find dy/dx. We write $F(x, y) = y^4 - 5y^2 - x^4 + 4x^2 = 0$ and compute

$$\frac{dy}{dx} = -\frac{\partial F/\partial x}{\partial F/\partial y} = -\frac{-4x^3 + 8x}{4y^3 - 10y} = \frac{2x^3 - 4x}{2y^3 - 5y}.$$

Then by the standard procedure we arrive at the points for which there are vertical tangents: $(\pm 2, 0)$ and $(0, 0)$; and for which there are horizontal tangents: $(0, \pm\sqrt{5/2})$ and $(0, 0)$. However, in checking these points, we notice that the point $(0, 0)$ actually makes the derivative take on the indeterminate form $0/0$. So how do we know that the tangents at the origin are really vertical (or horizontal)?

At this point, a newer textbook by a pair of eminent mathematicians instructs the students to use a calculator or computer. Using the calculator to find points on the curve close to the origin, they suggest calculating the average rate of change as an approximation to the slope of the tangent line. However, the old text by Love and Rainville [3] instructs students differently. That text includes the following definitions:

Definition. A *singular point* is a point on a curve at which the derivative becomes $0/0$. This indicates the presence of a *double point*—a point through which the curve passes twice. The tangent lines of the curve at a double point are found by considering only those terms of degree two in the equation and solving.

Since the terms of degree two are $-5y^2$ and $4x^2$, we ignore the other terms and solve $-5y^2 + 4x^2 = 0$ to get $y = \pm 2x/\sqrt{5}$; these are the tangent lines at the origin (which, by the way, proves that the tangents at the origin are neither vertical nor horizontal). This is a very nice idea![2] As to why this works, I will refer you to [3].

Another old textbook demonstrates how to apply the multivariable calculus method of optimization by Lagrange multipliers to our standard single-variable optimization problems (see [4]). The result is that a complex optimization problem turns into a simple of system of equations. These old textbooks are unique for other reasons, too. One of them being that there aren't too many problems. There seem to be just the right mix of routine and challenging problems for students to try—unlike modern textbooks. Which brings me to the last thing I learned:

Modern calculus textbooks have too many problems.

Open up any of the current favorite calculus textbooks and you will see something that should shock you: every section of every chapter has 70-100 problems. But my guess is that this doesn't shock you. This is probably something you are used to seeing in all calculus textbooks. But consider this: Is it possible for one student to do them all in a semester or two? Is it possible that any teacher, professor, or teaching assistant would grade them all even if a student did them? I do not think so. And the publishers of the textbooks don't think so either, or they wouldn't produce "teacher's guides" which suggest that students do "every third problem." If the publishers themselves produce guide books suggesting students are only to do every third problem, then why not eliminate two-thirds of the problems and cut the book down by 200 pages?

And this makes us lazy teachers. If the guidebook says to only assign every third problem, then that's what we do, isn't it? It's funny how we trust those guidebooks. Did we read every third problem to see if it was appropriate to gauge our students' understanding of what we taught them? Did we *do* every third problem in order to anticipate the struggles our students may face? More often than not, the answer is no—we simply make the students struggle with problems that may not have been appropriate and then wonder why they give up on doing homework as the semester drags on.

[2]This method only works if the double point is located at the origin. If the double point is located somewhere else, then one must translate the curve so that the double point lies on the origin. For instance, if the double point of the implicit curve $f(x, y) = 0$ is $(2, 3)$, then the curve must be translated to the left 2 units and down 3 units, creating the curve $f(x + 2, y + 3) = 0$.

Of course, the reason books have so many problems is that the publishers feel that to make a profit the book must be marketable to as many high schools, colleges, and universities as possible. So the publishers hire "assistants," "accuracy checkers," "art program designers," and myriad others to "help" the author write the book, develop problems, write solution manuals, and produce guidebooks which advocate doing every third problem. The textbook becomes a book written by committee instead of expressing the educational and mathematical vision of the author, and so the book becomes like all other textbooks: a 1200-page calculus doorstop. Why do we continue to purchase such nonsense?

References

[1] William Boyce and Richard DiPrima, *Calculus*, John Wiley & Sons, New York, 1988.

[2] Wilfred Kaplan, *Advanced Calculus*, 5th edition, Addison-Wesley, New York, 2002.

[3] Clyde Love and Earl Rainville, *Differential and Integral Calculus*, 5th edition, MacMillan, New York, 1954.

[4] Robert Osserman, *Two-Dimensional Calculus*, Harcourt Brace and World, New York, 1968.