What Is Wrong With the Definition of dy/dx?

Hugh Thurston

We shall use the notations dy/dx and $f'(x)$ freely and interchangeably. [1]

The fact that dy/dx and $f'(x)$ are not interchangeable is evident when you consider that one does not write $dy/d3$ for $f'(3)$. [2]

For a start, the definition is incomplete. It is always a good idea to know what we are talking about, but definitions of dy/dx do not say what the x and y are; in contrast, definitions of $f'(x)$ make it clear that f is a function and x a number.

Secondly, the definition is ambiguous. Most texts describe dy/dx as another notation for $f'(x)$ where $y = f(x)$. For this to be valid they should prove that if $f(x) = g(x)$ then $f'(x) = g'(x)$, but they don't. Indeed, it is hard to *prove* anything about dy/dx without knowing what x and y are. All we can say for certain is that they are not numbers: $f'(3)$ cannot be denoted by $dy/d3$.

Ambiguity breeds paradox. Whatever x may be, it is something that has values: in the familiar formula $dy/dx|_{x=c}$, c is a value of x. Moreover, y can be constant; we all know that if y is constant then dy/dx is zero. The formula

$$\frac{dx}{dy} = 1 \Big/ \frac{dy}{dx}$$

shows that x and y are the same kind of entity, so in principle x can be constant. Suppose that x is constant with value 1. Then $x^2 = x^3$, $dx^2/dx = dx^3/dx$, and $2x = 3x^2$. But $2x$ has value 2 and $3x^2$ has value 3. The obvious objection is that it is nonsense to differentiate with respect to a constant—we cannot have a rate of change with respect to something that is not changing. This objection may be obvious, but it is not valid; if x cannot be constant in dy/dx there should be something in the definition that implies this.

So what are x and y? We can take a hint from the fact that they have values. There is a familiar entity that has values—the function. Or we can make the reasonable suggestion that the x and y in dy/dx are the same as in dx and dy. In the modern (Fréchet) theory of differentials, x and y are functions. We don't apply Fréchet theory to elementary calculus, but if we did the x and y would be the familiar type of function whose values and arguments are real numbers. From now on by "function" I shall mean this type of function.

Now what is dy/dx? First, look to Leibniz. His dx was an infinitesimal increment in x. An increment in x is $x(c+h) - x(c)$ and the corresponding increment in y is $y(c+h) - y(c)$. Then the value of dy/dx at c is

$$\frac{y(c+h) - y(c)}{x(c+h) - x(c)} \tag{1}$$

where h is infinitesimal. Nowadays instead of having h infinitesimal we have it approach zero. Then (1) becomes $y'(c)/x'(c)$ if x and y are differentiable at c and $x'(c) \neq 0$.

We find the same result if we consider tangents: the slope of a secant of the graph of y against x is (1), and the slope of the tangent is its limit. Rates of change lead to the same result: if x and y represent two quantities that vary with time the average rate of change of the second quantity with respect to the first between times c and $c + h$ is (1) and the instantaneous rate of change is its limit.

All this suggests the following definition.

Definition. If x and y are functions, dy/dx is y'/x'.

From this definition we can prove the familiar rules of differentiation, no longer as mere rules but as properly-stated theorems. It is convenient to use the not-uncommon figures of speech "f exists at c" for "$f(c)$ exists" and "$f = g$ at c" for "$f(c) = g(c)$".

For example

Theorem (chain rule). *If x, y and z are functions,*

$$\frac{dz}{dx} = \frac{dz}{dy}\frac{dy}{dx}$$

wherever the right-hand side exists.

Proof: If the right-hand side exists at c, then

$$\left(\frac{dz}{dy}\frac{dy}{dx}\right)(c) = \frac{dz}{dy}(c)\frac{dy}{dx}(c) = \frac{z'(c)y'(c)}{y'(c)x'(c)} = \frac{z'(c)}{x'(c)} = \frac{dz}{dx}(c).$$

We can also say clearly and definitely:

Theorem. *If x is constant, dy/dx does not exist anywhere.*

Proof: x' has the value 0.

Theorem. *If x is increasing on an interval I and dy/dx is negative on I then y is decreasing on I.*

There are analogous results if x is decreasing or dy/dx is positive or both. The proofs are obvious.

Our definition legitimizes the use of parametric and implicit differentiation. For example,

$$x(t) = 2\cos t,\ y(t) = \sin t$$

is a parametrization of the ellipse $x^2 + 4y^2 = 4$. We have

$$\frac{dy}{dx}(t) = \frac{y'(t)}{x'(t)} = -\frac{\cos t}{2\sin t},$$

giving the slope at $(2\cos t, \sin t)$. This calculation, and the corresponding implicit one, occur in texts, but in any text which defines dy/dx only where y is a function of x they are necessarily invalid.

Finally how does our definition relate to the traditional one? First, what is $f(x)$ if x is a function? It is the function $t \to f(x(t))$. For example, the speed of a body moving with positive acceleration is a function of the distance covered: if y denotes the speed and x the distance, then $y = f(x)$ for some f. (If the acceleration a is constant, $f(x) = \surd(2ax)$.) At time t, when the distance is $x(t)$, the speed is $f(x(t))$, so that $y(t) = f(x(t))$.

It follows that if $y = f(x)$ then $y' = f'(x)x'$ wherever the right-hand side exists, and so $dy/dx = f'(x)$ wherever the right-hand side exists and x' has a non-zero value.

Adopting the definition suggested here would not alter the well-known and universally-accepted formulas involving the Leibnizian derivative, but it would give them a sound basis.

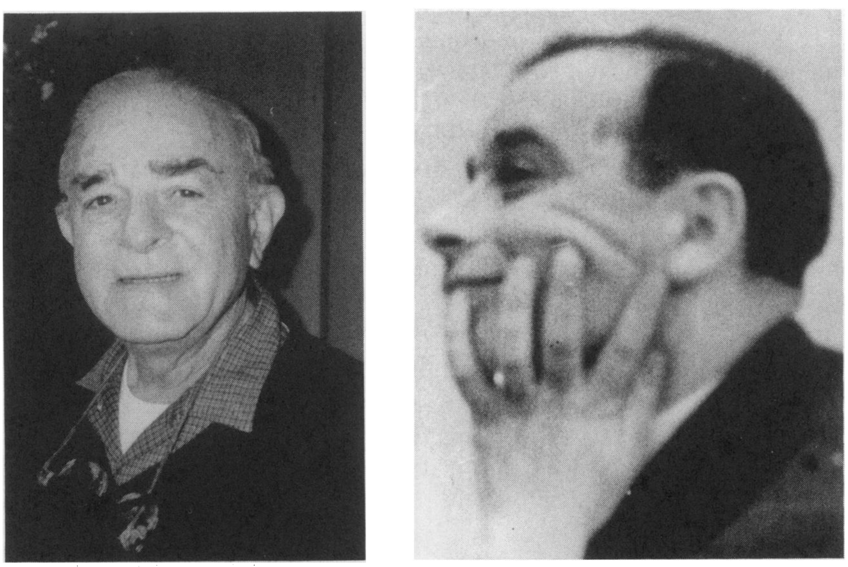
REFERENCES

1. Serge Lang, *A First Course in the Calculus* (third edition) 1971, p. 248.
2. A. R. Pargeter, *Mathematical Gazette*, 54 (1970) *p*. 165.

Department of Mathematics
University of British Columbia
#121-1984 Mathematics Road
Vancouver, B.C., Canada V6T 1Y4

PICTURE PUZZLE

(from the collection of Paul Halmos)



They frequently collaborated, but these photos were taken far apart: the first in 1986 and the second in 1938.

(see page 923.)