Understanding the "Chain Rule" by Deriving Your Own Version

July 29, 2018

James Smith
nitac14b@yahoo.com
https://mx.linkedin.com/in/james-smith-1b195047

Abstract

Because the Chain Rule can confuse students as much as it helps them solve real problems, we put ourselves in the shoes of the mathematicians who derived it, so that students may understand the motivation for the rule; its limitations; and why textbooks present it in its customary form. We begin by finding the derivative of $\sin 2x$ without using the Chain Rule. That exercise, having shown that even a comparatively simple compound function can be bothersome to differentiate using the definition of the derivative as a limit, provides the motivation for developing our own formula for the derivative of the general compound function $g\left[f\left(x\right)\right]$. In the course of that development, we see why the function f must be continuous at any value of x to which the formula is applied. We finish by comparing our formula to that which is commonly given.

Contents

1 Introduction

When we are learning a new technique in mathematics, we benefit from familiarizing ourselves with the type of problem that the method was developed to solve. We also benefit from struggling with a few problems of that sort before being shown the technique in its modern form. In that way, we are better prepared to understand that version of that technique, as well as its derivation. We also become better problem-solvers in general.

In this document, we will develop our own formula for the derivative of a composite function, then compare it to a version of the Chain Rule that is found in many standard calculus textbooks. As a motivational example (that is, to help us see why some sort of Chain Rule would be desirable), we'll begin by finding $\frac{\mathrm{d}\sin 2x}{\mathrm{d}x}$ using the definition

$$\left.\frac{\mathrm{d}u\left(x\right)}{\mathrm{d}x}\right|_{x=a}=\lim_{\delta\to 0}\frac{u\left(a+\delta\right)-u\left(a\right)}{\delta}, \tag{1.1}$$

followed by the Law of Universal Generalization.

2 Motivational Example: Derivative of $\sin 2x$ from the Definition of the Derivative as a Limit

For the real number a, arbitrary, the definition in Eq. (1.1) gives

$$\left.\frac{\mathrm{d}\sin 2x}{\mathrm{d}x}\right|_{x=a}=\lim_{\delta\to 0}\frac{\sin 2\left(a+\delta\right)-\sin 2a}{\delta}.$$

The next several steps use trigonometric identities for sums and doubles of angles to transform the right-hand side.

The trig identities that we need:
(1) $\sin 2\theta = 2\sin\theta\cos\theta$

(2) $\sin(\alpha+\beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta$

(3) $\cos(\alpha+\beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta$.

$$\begin{aligned}
\left.\frac{\mathrm{d}\sin 2x}{\mathrm{d}x}\right|_{x=a} &= \lim_{\delta\to 0}\frac{2\sin(a+\delta)\cos(a+\delta)-\sin 2a}{\delta} \\
&= \lim_{\delta\to 0}\frac{2\left[\sin(a+\delta)\right]\left[\cos(a+\delta)\right]-\sin 2a}{\delta} \\
&= \lim_{\delta\to 0}\frac{2\left[\sin a\cos\delta+\cos a\sin\delta\right]\left[\cos a\cos\delta-\sin a\sin\delta\right]-\sin 2a}{\delta} \\
&= \lim_{\delta\to 0}\frac{2\sin a\cos a\left(\cos^2\delta-\sin^2\delta-1\right)+2\left(\cos^2 a-\sin^2 a\right)\sin\delta\cos\delta}{\delta} \\
&= \lim_{\delta\to 0}\frac{2\sin a\cos a\left(-2\sin^2\delta\right)+2\cos 2a\sin\delta\cos\delta}{\delta}.
\end{aligned}$$

Next, we transform the right-hand side in a way that will enables to use theorems about limits.

$$\left.\frac{\mathrm{d}\sin 2x}{\mathrm{d}x}\right|_{x=a} = \lim_{\delta\to 0}\left[\frac{-4\sin a\cos a\left(\sin^2\delta\right)}{\delta}+\frac{2\cos 2a\sin\delta\cos\delta}{\delta}\right]$$

Now, we'll use those theorems about limits.

$$\begin{aligned}
\left.\frac{\mathrm{d}\sin 2x}{\mathrm{d}x}\right|_{x=a} &= \lim_{\delta\to 0}\frac{-4\sin a\cos a\left(\sin^2\delta\right)}{\delta}+\lim_{\delta\to 0}\frac{2\cos 2a\sin\delta\cos\delta}{\delta} \\
&= -4\sin a\cos a\lim_{\delta\to 0}\left[\left(\frac{\sin\delta}{\delta}\right)\sin\delta\right]+2\cos 2a\lim_{\delta\to 0}\left[\left(\frac{\sin\delta}{\delta}\right)\cos\delta\right] \\
&= -4\sin a\cos a\underbrace{\left[\lim_{\delta\to 0}\frac{\sin\delta}{\delta}\right]}_{=1}\underbrace{\left[\lim_{\delta\to 0}\sin\delta\right]}_{=0}+2\cos 2a\underbrace{\left[\lim_{\delta\to 0}\frac{\sin\delta}{\delta}\right]}_{=1}\underbrace{\left[\lim_{\delta\to 0}\cos\delta\right]}_{=1} \\
&= 2\cos 2a.
\end{aligned}$$

We conclude by saying that because a was an arbitrary real number, the result is valid for all real numbers. Customarily, we communicate that conclusion by writing

$$\frac{\mathrm{d}\sin 2x}{\mathrm{d}x} = 2\cos 2x.$$

WOW! That was a lot of work to find the derivative of such a simple function. We might well dread trying to find the derivative of, say, $\sin\sqrt{1+\log x}$ by the same route. That is, by starting from the definition in (Eq. (1.1)). Let's see if we can find a better idea.

3 Looking for an Easier Route

Anyone who has solved math problems "by hand" or with spreadsheets has seen several benefits of treating functions like $\sin 2x$ as composites of the form

$v\left[u\left(x\right)\right]$. (In the case of $\sin 2x$, $u\left(x\right)=2x$, and v is the sine function.) Therefore, why not attempt to derive a formula for the derivative of the generic composite function $g\left[f\left(x\right)\right]$? We want our formula to be applicable to as many types of functions as possible, so we'll accept restrictions upon g, f, and their domains only when necessary. We'll begin by writing, for $x=a$, arbitrary,

$$\left.\frac{\mathrm{d}g\left[f\left(x\right)\right]}{\mathrm{d}x}\right|_{x=a}=\lim_{\delta\to 0}\frac{g\left[f\left(a+\delta\right)\right]-g\left[f\left(a\right)\right]}{\delta}. \tag{3.1}$$

The expression on the right-hand side appears unhelpful, so we'll look for ideas that might suggest ways to transform it. We're searching for notions, so for now we won't pay much attention to rigor—time for that later. If we bear in mind that $\frac{\mathrm{d}g\left[f\left(x\right)\right]}{\mathrm{d}x}$ is 'the rate of change of g with respect to x", we might jot down (informally)

$$\boxed{\text{Rate of change of } g \text{ with respect to x}}=\left\{\boxed{\text{Rate of change of } f \text{ with respect to } x}\right\}\left\{\boxed{\text{Rate of change of } g \text{ with respect to } f}\right\}.$$

Continuing to think informally, might rewrite that note as

$$\frac{\mathrm{d}g}{\mathrm{d}x}=\frac{\mathrm{d}f}{\mathrm{d}x}\frac{\mathrm{d}g}{\mathrm{d}f}. \tag{3.2}$$

This idea has intuitive appeal. Let's test it on our result for $\frac{\mathrm{d}\sin 2x}{\mathrm{d}x}$. As we noted above, our "f" in that case is $2x$, so $\frac{\mathrm{d}f}{\mathrm{d}x}$ would be 2. Ou "g" is the sine function. Viewing $2x$ as a single variable, the derivative of $\sin 2x$ with respect to $2x$ would be $\cos 2x$. Thus, $\frac{\mathrm{d}g}{\mathrm{d}f}$ in our case would be $\cos 2x$. Putting these ideas together,

$$\begin{aligned}\frac{\mathrm{d}\sin 2x}{\mathrm{d}x}&=\frac{\mathrm{d}f}{\mathrm{d}x}\frac{\mathrm{d}g}{\mathrm{d}f}\\&=\left[2\right]\left[\cos 2x\right]\\&=2\cos 2x,\end{aligned}$$

which is the result that we we obtained in our motivational example.

Although we would appear to be on the right track, we can't trust our idea that $\frac{\mathrm{d}g}{\mathrm{d}x}=\frac{\mathrm{d}f}{\mathrm{d}x}\frac{\mathrm{d}g}{\mathrm{d}f}$ without deriving it rigorously—for example, from Ec. (3.1))—and expressing it clearly. How might we do that? Recognizing that Ec. (3.1) refers specifically to the value of the derivative for $x = a$, we might write

$$\boxed{\begin{array}{l}\text{Rate of}\\ \text{change of } g\\ \text{with respect}\\ \text{to x}\end{array}} = \left\{ \left.\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right|_{x=a} \right\} \left\{ \boxed{\begin{array}{l}\text{Rate of change}\\ \text{of } g \text{ with re-}\\ \text{spect to } f \text{ at}\\ x = a\end{array}} \right\}, \text{ and therefore}$$

$$\left.\frac{\mathrm{d}g[f(x)]}{\mathrm{d}x}\right|_{x=a} = \left\{ \lim_{\delta\to 0} \frac{f(a+\delta)-(a)}{\delta} \right\} \left\{ \boxed{\begin{array}{l}\text{Rate of change}\\ \text{of } g \text{ with re-}\\ \text{spect to } f \text{ at}\\ x = a\end{array}} \right\}. \tag{3.3}$$

At this point, we might note that the quantity in the box on the right-hand side is a limit of "something" as $\delta \to 0$:

$$\left.\frac{\mathrm{d}g[f(x)]}{\mathrm{d}x}\right|_{x=a} = \left\{ \lim_{\delta\to 0} \frac{f(a+\delta)-(a)}{\delta} \right\} \left\{ \lim_{\delta\to 0} \left[\text{“}\textit{Something}\text{”}\right] \right\}. \tag{3.4}$$

But what is that "Something"? Comparing the right-hand sides of Ecs. (3.1) and (3.4), and using the theorem that "the limit of a product of functions is the product of the functions' limits", we reason as follows:

$$\begin{aligned} \left.\frac{\mathrm{d}g[f(x)]}{\mathrm{d}x}\right|_{x=a} &= \left.\frac{\mathrm{d}g[f(x)]}{\mathrm{d}x}\right|_{x=a} \\ \left\{ \lim_{\delta\to 0} \frac{f(a+\delta)-(a)}{\delta} \right\} \left\{ \lim_{\delta\to 0} \left[\text{“}\textit{Something}\text{”}\right] \right\} &= \lim_{\delta\to 0} \frac{g[f(a+\delta)]-g[f(a)]}{\delta} \\ \lim_{\delta\to 0} \left\{ \left[\frac{f(a+\delta)-(a)}{\delta}\right] \left[\text{“}\textit{Something}\text{”}\right] \right\} &= \lim_{\delta\to 0} \frac{g[f(a+\delta)]-g[f(a)]}{\delta} \\ \therefore \quad \text{“}\textit{Something}\text{”} &= \frac{g[f(a+\delta)]-g[f(a)]}{f(a+\delta)-f(a)}. \end{aligned}$$

Now, we can return to Ec. (3.4) to write

$$\begin{aligned} \left.\frac{\mathrm{d}g[f(x)]}{\mathrm{d}x}\right|_{x=a} &= \left\{ \lim_{\delta\to 0} \frac{f(a+\delta)-(a)}{\delta} \right\} \left\{ \lim_{\delta\to 0} \frac{g[f(a+\delta)]-g[f(a)]}{f(a+\delta)-f(a)} \right\} \\ &= \left\{ \left.\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right|_{x=a} \right\} \left\{ \lim_{\delta\to 0} \frac{g[f(a+\delta)]-g[f(a)]}{f(a+\delta)-f(a)} \right\}. \end{aligned} \tag{3.5}$$

Having written that result, we need to recall that it is true only if the derivative of f with respect to x exists at a.

A restriction upon the formula that we're attempting to develop: the derivative of f with respect to x must exist at $x = a$.

We've just placed our first restriction upon whatever result we may obtain from our derivation.

Our question now is what to to with the remaining limit on the right-hand side of Ec. (3.5):

$$\lim_{\delta\to 0} \frac{g[f(a+\delta)]-g[f(a)]}{f(a+\delta)-f(a)}$$

That's quite a "busy" expression, so let's draw a graph to help us "get our minds around it". We'll start with a graph of $f(x)$ (Fig. 1).

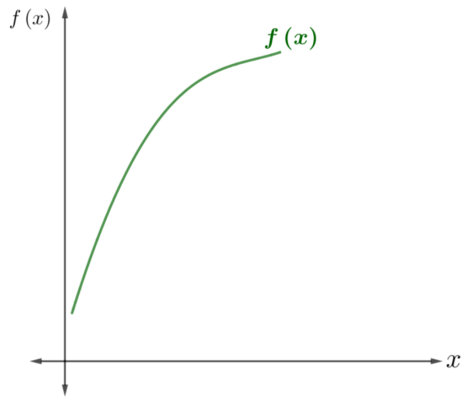

Figure 1: Our first step in constructing a graph that might help us to understand the limit on the right-hand side of Eq. (3.5): the graph of $f(x)$.

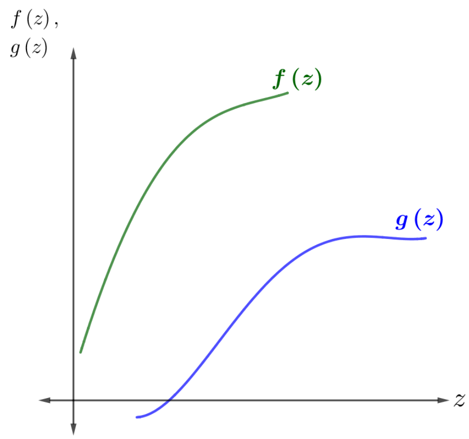

Figure 2: To eliminate a possible confusion, we've graphed both f and g as functions of the real variable z.

Next, we'll want to add the graph of $g[f(x)]$. But how do we do that, on a graph whose horizontal axis is x?

At this point, we might realize that we've been a little careless with our use of symbols. We're accustomed to using the single symbol "x" both to represent the independent variable in a problem (as we have here with $f(x)$), and as a coordinate along the real-number line (as in our graph). That dual meaning seldom causes trouble for us, but now it has.

"Defined" means that for every real number b, there exists a unique real number $f(b)$. Similarly for g.

To find a way forward, let's consider the case where both f and g are defined for every real number z. (We'll discuss more-complicated cases later.) We'll graph both functions in that way (Fig. 2). Now, along the horizontal axis, we'll locate the point for the number a at which we're evaluating our derivative $\frac{\mathrm{d}g[f(x)]}{\mathrm{d}x}$. On the vertical axis, we'll locate the point for the number $f(a)$ (Fig. 3).

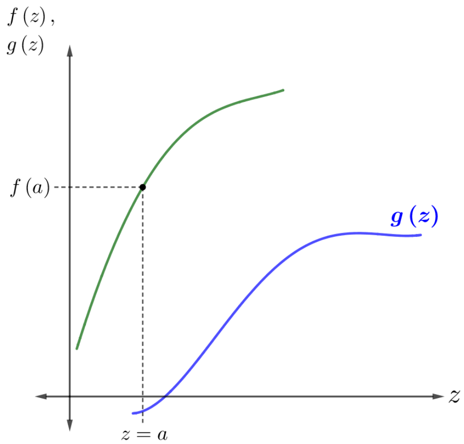

Figure 3: Along the horizontal axis, we've located the point for the number a at which we're evaluating our derivative $\dfrac{\mathrm{d}g\left[f(x)\right]}{\mathrm{d}x}$. On the vertical axis, we've located the point for the number $f(a)$.

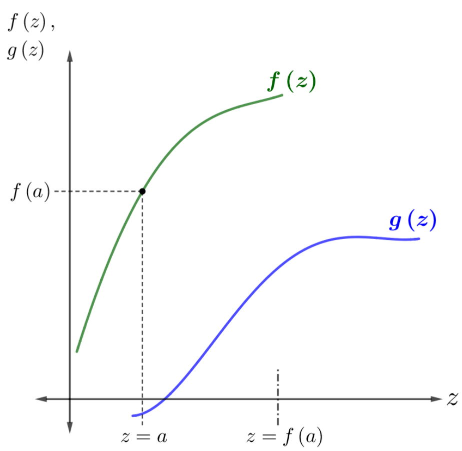

Figure 4: After locating the point on the horizontal axis for the number $f(a)$.

Because the function g is defined for every real number, it's defined for the specific real number $f(a)$. Therefore, our next step is to locate the point on the horizontal axis for that number (Fig. 4). The value of g, evaluated at $z = f(a)$, is some specific real number that we'll write as $g\left[f(a)\right]$. In Fig. 5, we locate the point for that number along the vertical axis.

We seem to be progressing, but we've yet to incorporate δ. We should know how to do that; we did it many times in our first classes on derivatives as limits. Still, before adding to our graph the points that involve δ, we want to think a bit about our goal: we want to understand what happens when δ approaches zero. To that end, we first attempt to understand the situation that exists when δ is some suitably small, non-zero number. Taking δ as positive for the time being (Fig. 6), we locate the point for the number $z = a + \delta$ on the horizontal

The set-theoretic concept of a function may be helpful here: The function f is the set of ordered pairs (b, c) such that no two pairs have b as their first element. We can call c "the value of f at $z = b$". When making a graph of f, we "highlight" those points whose horizontal coordinate is the first element of some pair, and whose vertical coordinate is the second element of that same pair.

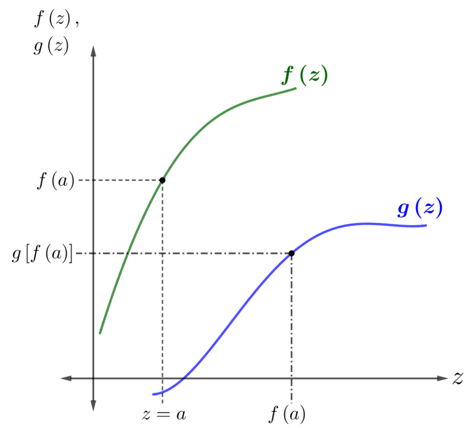

Figure 5: After locating the point on the vertical axis for the number $g[f(a)]$.

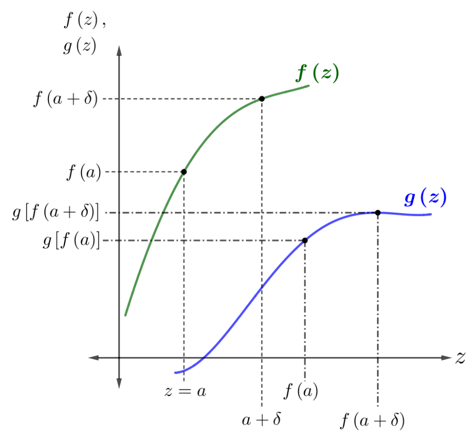

Figure 6: After locating the point for $g[(a+\delta)]$.

axis, and that for the number $f(a+\delta)$ on the vertical axis. Then, we locate $f(a+\delta)$ on the horizontal axis, and $g[f(a+\delta)]$ on the vertical axis.

We're ready, finally, to investigate $\lim_{\delta\to 0}\dfrac{g[f(a+\delta)]-g[f(a)]}{f(a+\delta)-f(a)}$. To avoid distractions, we'll eliminate the portions of our graph that don't concern g (Fig. 7). We'll also draw a straight line connecting the indicated points on the curve for $g(z)$.

We can let our early experiences with derivatives as limits guide us now. As the interval between $z=f(a)$ and $z=f(a+\delta)$ shrinks, the straight line that we drew becomes the line tangent to the graph of g at $z=f(a)$ (Fig. 8). The slope of that tangent line is $\left.\dfrac{\mathrm{d}g(z)}{\mathrm{d}z}\right|_{z=f(a)}$. Therefore, if $f(a+\delta)-f(a)$ goes to zero as δ itself shrinks to zero, then

$$\lim_{\delta\to 0}\frac{g[f(a+\delta)]-g[f(a)]}{f(a+\delta)-f(a)}=\left.\frac{\mathrm{d}g(z)}{\mathrm{d}z}\right|_{z=f(a)}. \tag{3.6}$$

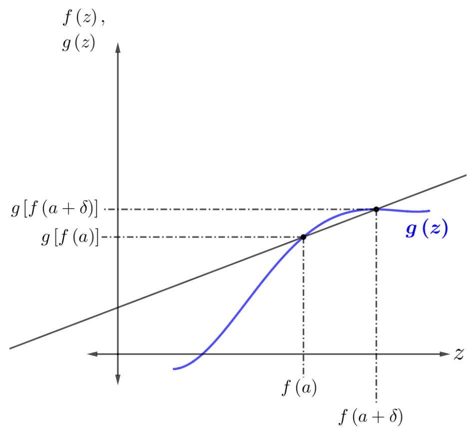

Figure 7: Focusing on the curve for g. We've added the secant line as preparation for considering what occurs when $\delta \to 0$.

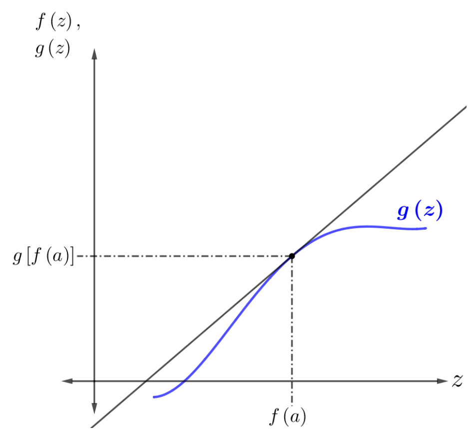

Figure 8: The tangent that the secant line shown in Fig. 7 approaches if $f(a+\delta) \to f(a)$ as $\delta \to 0$. The slope of the tangent is $\left.\frac{\mathrm{d}g(z)}{\mathrm{d}z}\right|_{z=f(a)}$.

But notice the "if": as we know, not all functions behave as stated for every real number.

Where does that realization leave us? The bad news is that if we're dealing with a function f and a number a such that $f(a+\delta)$ does not go to $f(a)$ as δ goes to zero, then we're stuck: nothing can be done. The good news is that many common functions do have the required behavior: they're the type that mathematicians call continuous. That is, the functions $u(z)$ such that for every real number c,

$$\lim_{z \to c^+} u(z) = \lim_{z \to c^-} u(z) = u(c).$$

Polynomials and $\sin x$ are examples of continuous functions.

Functions can also be *piecewise continuous*; that is, continuous on certain intervals. The same arguments that we're using here work for that type of function as well.

The need for $f(z)$ to be continuous at a becomes apparent when f has the behavior shown in Fig. 9. In such a case, the point on the horizontal axis for

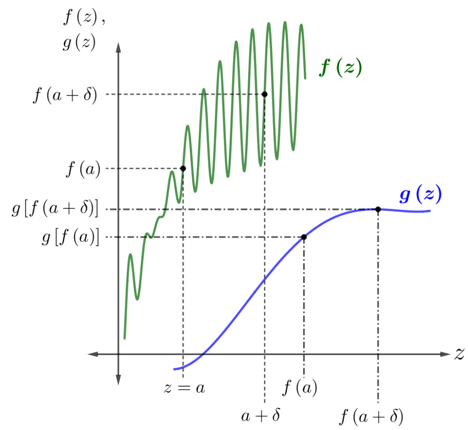

Figure 9: A function in which $f(a+\delta)$ would be equal to $\fallingdotseq(a)$ several times as $\delta \to 0$, at each of which the denominator in Eq. (3.6) would be zero. Those subtleties require treatment that is beyond the scope of this document.

$f(a+\delta)$ will alternate between being to the left and the right of that for $f(a)$. Nevertheless, the length of the interval will shrink to zero as δ itself goes to zero. Note that for some values of δ in Fig. 9, $f(a+\delta) = f(a)$, making the denominator zero in the limit in Eq. (3.6). Those subtleties require treatment that is beyond the scope of this document.

However, those considerations should not distract us from what we've accomplished: accepting the restriction that f must be continuous at a is a small price to pay for being able to reduce a monstrosity like $\lim_{\delta\to 0} \dfrac{g[f(a+\delta)] - g[f(a)]}{f(a+\delta) - f(a)}$ to $\left.\dfrac{\mathrm{d}g(z)}{\mathrm{d}z}\right|_{z=f(a)}$. Using that result, we can write that if f is continuous at a, and if $\dfrac{\mathrm{d}g(z)}{\mathrm{d}z}$ is continuous at $f(a)$, then

$$\left.\frac{\mathrm{d}g[f(a)]}{\mathrm{d}x}\right|_{x=a} = \left[\left.\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right|_{x=a}\right]\left[\left.\frac{\mathrm{d}g(z)}{\mathrm{d}z}\right|_{z=f(a)}\right]. \tag{3.7}$$

The right-hand side of that equation is messy because of the (apparently) different variables x and z. To clean it up, we can ask ourselves what those variables mean in this context. We'll start with $\left.\dfrac{\mathrm{d}f(x)}{\mathrm{d}x}\right|_{x=a}$: that expression means "the rate of change of the dependent variable f with respect to its independent variable, when the value of the latter is a". In the context of our present problem, x and z refer to the same variable. Therefore, we'll use z, and rewrite Eq. (3.7) as

$$\left.\frac{\mathrm{d}g[f(a)]}{\mathrm{d}x}\right|_{x=a} = \left[\left.\frac{\mathrm{d}f(z)}{\mathrm{d}z}\right|_{z=a}\right]\left[\left.\frac{\mathrm{d}g(z)}{\mathrm{d}z}\right|_{z=f(a)}\right]. \tag{3.8}$$

Let's test that result on the function $\sin 2x$, which we used in our motivational example. Our $f(x)$ is $2x$, and our g is the sine function. The procedure is shown in Table 1 to find the derivative of $\sin 2x$ according to Eq. (3.8).

Table 1: Implementation of Eq. (3.8)

Step	Implementation for $g\left[f\left(x\right)\right]=\sin 2x$
Identify $f\left(z\right)$ and $g\left(z\right)$	$f\left(z\right)=2z,\ g\left(z\right)=\sin z$
Identify formulas for $\dfrac{\mathrm{d}f\left(z\right)}{\mathrm{d}z}$, and $\dfrac{\mathrm{d}g\left(z\right)}{\mathrm{d}z}$	$\dfrac{\mathrm{d}2z}{\mathrm{d}z}=2$, and $\dfrac{\mathrm{d}\sin z}{\mathrm{d}z}=\cos z$
Evaluate $\dfrac{\mathrm{d}f\left(z\right)}{\mathrm{d}z}$ at $z=a$, and $\dfrac{\mathrm{d}g\left(z\right)}{\mathrm{d}z}$ at $z=f\left(a\right)$	$2\|_{z=a}=2$; and $\cos z\ \|_{z=2a}=\cos 2a$
$\left.\dfrac{\mathrm{d}g\left[f\left(a\right)\right]}{\mathrm{d}z}\right\|_{x=a}=\left[\left.\dfrac{\mathrm{d}f\left(z\right)}{\mathrm{d}z}\right\|_{z=a}\right]\left[\left.\dfrac{\mathrm{d}g\left(z\right)}{\mathrm{d}z}\right\|_{z=f\left(a\right)}\right]$	$\left.\dfrac{\mathrm{d}g\left[\sin 2x\right]}{\mathrm{d}z}\right\|_{x=a}=\left[2\right]\left[\cos 2a\right]$

Because the procedure worked, we now invoke the Law of Universal Generalization to write that because (1) the function $f=2z$ is continuous for all values of z, (2) $\dfrac{\mathrm{d}2z}{\mathrm{d}z}$ exists at all z, and (3) $\cos z$ exists at all z,

$$\frac{\mathrm{d}\sin 2x}{\mathrm{d}x}=2\cos 2x.$$

Now that we're sure our method for finding the derivative of a compound function is sound, we'll want to compare our method to the standard formulation of the Chain Rule.

4 Comparison with the Usual Form of the Chain Rule

A typical presentation of the Chain Rule is

> If a variable g depends on the variable f, which itself depends on the variable x, so that f and g are therefore dependent variables, then g, via the intermediate variable of f, depends on x as well. The chain rule then states [1]
>
> $$\frac{\mathrm{d}g}{\mathrm{d}x}=\frac{\mathrm{d}g}{\mathrm{d}f}\frac{\mathrm{d}f}{\mathrm{d}x}. \tag{4.1}$$

[1]Paraphrase of Wikipedia's article "Chain rule", accessed 22 Julio 2018.

Eq. (4.1) is identical to our "intuitive" Eq. (3.2). In both, the derivative $\frac{\mathrm{d}f}{\mathrm{d}x}$ is the customary way of writing the generalization of our $\frac{\mathrm{d}f\,(z)}{\mathrm{d}z}\,|_{z=a}$ (in Eq. (3.8)) to the whole set of real numbers. (Or more accurately, to those at which $\frac{\mathrm{d}f}{\mathrm{d}z}$ exists.) But what about the factor $\frac{\mathrm{d}g}{\mathrm{d}f}$ in Eqs. (3.2) and (4.1)? It must be equal to the factor $\frac{\mathrm{d}g\,(z)}{\mathrm{d}z}\,|_{z=f(a)}$ in our Eq. (3.8). Can we establish that equality rigorously?

Let's review the analysis through which we established that

$$\lim_{\delta\to 0}\frac{g\,[f\,(a+\delta)]-g\,[f\,(a)]}{f\,(a+\delta)-f\,(a)}=\left.\frac{\mathrm{d}g\,(z)}{\mathrm{d}z}\right|_{z=f(a)},$$

with the restriction that f must be continuous at a. We accepted that restriction because it ensured that $[f\,(a+\delta)-f\,(a)]\to 0$ as $\delta\to 0$. Therefore, the expressions

For any given f, g, and a for which the limit $\lim_{\delta\to 0}\frac{g\,[f\,(a+\delta)]-g\,[f\,(a)]}{f\,(a+\delta)-f\,(a)}$ exists, that limit is a specific real number.

$$\lim_{\delta\to 0}\frac{g\,[f\,(a+\delta)]-g\,[f\,(a)]}{f\,(a+\delta)-f\,(a)}$$

and

$$\lim_{[f(a+\delta)-f(a)]\to 0}\frac{g\,[f\,(a+\delta)]-g\,[f\,(a)]}{f\,(a+\delta)-f\,(a)}$$

are the same number. The latter expression is the definition, given in the form of a limit, of $\frac{\mathrm{d}g}{\mathrm{d}f}\,|_{f=f(a)}$. We'd generalize that result by writing simply $\frac{\mathrm{d}g}{\mathrm{d}f}$.

Thus, $\frac{\mathrm{d}g}{\mathrm{d}f}$ in Eqs. (3.2) and (4.1) is indeed the generalization of the factor $\left.\frac{\mathrm{d}g\,(z)}{\mathrm{d}z}\right|_{z=f(a)}$ that we identified in our own version of the Chain Rule.