Assignment 2 Problem 7
Third Partial Solution

John McCuan

February 6, 2022

Here is a third proof of the chain rule for a composition of holomorphic functions. This proof is, in a certain sense, "the same" as the second proof, but there are some new elements. In particular, I'm going to discuss "little O" and "big O" notation[1] in more detail and give a lemma concerning those concepts. I think the resulting proof will now be under a page. I'll emphasize once again that this is really just calculus rather than complex analysis.

1 Little O and Big O

I have mentioned that a (complex valued[2]) function F of a (complex) variable h is said to be "little O" of h which we write as

$$F(h) = \circ(h)$$

if

$$\lim_{h\to 0} \frac{F(h)}{h} = 0.$$

In terms of tolerances ϵ and δ, this means for each $\epsilon > 0$, there is some $\delta > 0$ such that

$$\left|\frac{F(h)}{h}\right| < \epsilon \qquad \text{whenever} \qquad 0 < |h| < \delta. \tag{1}$$

Naturally, one should undertake to complete the following:

Exercise 1 *A function $f : \Omega \to \mathbb{C}$ defined on an open set $\Omega \subset \mathbb{C}$ is differentiable at $z \in \Omega$ if and only if there exists a complex number w for which*

$$f(z+h) - f(z) - wh = \circ(h).$$

There is another similar notation called "big O" notation. A function $F = F(h)$ as above is "big O" of h, written

$$F(h) = O(h),$$

if

$$\limsup_{h\to 0} \left|\frac{F(h)}{h}\right| < \infty.$$

In technical terms, this means the quotient $F(h)/h$ remains bounded as h tends to zero, or there is some $M > 0$ and some $\delta > 0$ such that

$$\left|\frac{F(h)}{h}\right| < M \qquad \text{whenever} \qquad 0 < |h| < \delta.$$

[1] Incidentally, this notation is also known as Landau's notation.

[2] More generally, I guess one can formulate this kind of definition for a function taking values in a normed space.

In this form, the "big O" condition looks relatively closely related to the "little O" condition expressed in (1).

Here is a result about products we can use to encapsulate some of the grimy details in the proof of the chain rule.

Lemma 1 *If*

(i) $G(\mu) = \circ(\mu)$,

(ii) $G(0) = 0$, *and*

(iii) $F(h) = O(h)$,

then

$$G(F(h)) = \circ(h).$$

Remark: It is quite natural to ask at this point, in view of the statement of the lemma, about the domains of the functions involved. The natural domain for the functions F in the definitions of "big O" and "little O" is **any punctured neighborhood of** $0 \in \mathbb{C}$. The evaluation $G(0)$ in the statement of the lemma implicitly requires a full neighborhood of $0 \in \mathbb{C}$ in the domain of G. The composition $G \circ F$, moreover, requires the existence of some particular neighborhood U_0 of $0 \in \mathbb{C}$ in the domain of G for which the values of F satisfy

$$F(h) \in U_0 \qquad \text{at least for some neighborhood } 0 < |h| < \delta_0.$$

Proof of Lemma 1: There is some $M > 0$ and some $\delta_1 > 0$ such that $0 < |h| < \delta_1$ implies

$$\left| \frac{F(h)}{h} \right| < M. \tag{2}$$

Let $\epsilon > 0$.

There is some δ_2 for which $0 < |\mu| < \delta_2$ implies

$$\left| \frac{G(\mu)}{\mu} \right| < \frac{\epsilon}{M}. \tag{3}$$

Now, consider $H(h) = G \circ F(h)$ and $h \in \mathbb{C} \backslash \{0\}$. If $F(h) \neq 0$, we can write

$$\left| \frac{H(h)}{h} \right| = \left| \frac{G(F(h))}{F(h)} \right| \; \left| \frac{F(h)}{h} \right|. \tag{4}$$

Notice that for

$$0 < |h| < \delta = \min\left\{ \delta_1, \frac{\delta_2}{M} \right\}, \tag{5}$$

we know from (2) that $|F(h)| < |h|M < \delta_2$, and we get from (4) and (3), still under the assumption $F(h) \neq 0$,

$$\left| \frac{H(h)}{h} \right| < \frac{\epsilon}{M} \, M = \epsilon.$$

In the complementary situation $F(h) = 0$, we see from condition **(ii)** of the lemma that

$$\frac{H(h)}{h} = \frac{G(F(h))}{h} = 0.$$

We conclude that under the specification (5) we must have

$$\left|\frac{H(h)}{h}\right| < \epsilon.$$

This is what is required to show $H(h) = G \circ F(h) = \circ(h)$. $\square$

I hope some elements of my second proof of the chain rule are isolated and evident in this lemma. Note that condition **(ii)** can not be omitted. For this reason, I think it was probably a bit of an overstatement (or oversight) of Stein to say this result follows "easily" from the "little O" formulation of differentiability.

2 Proof of the Chain Rule

Notice there are sort of a lot of conditions in Lemma 1. This accounts for some of the (necessary) complication in my previous proofs of the chain rule. Here is the revised proof:

We wish to show

$$(g \circ f)(z+h) - (g \circ f)(z) - (g' \circ f)(z)\ f'(z)\ h = \circ(h). \tag{6}$$

We set

$$G(\mu) = g(f(z)+\mu) - g(f(z)) - g'(f(z))\ \mu$$

and

$$k(h) = f(z+h) - f(z).$$

Notice that by the definition of differentiability of g at $f(z)$, we have

$$G(\mu) = \circ(\mu)$$

in accordance with hypothesis **(i)** of the lemma. Furthermore, $0 \in \mathbb{C}$ is in the domain of G with $G(0) = 0$ so that hypothesis **(ii)** of the lemma holds as well.

By the differentiability of f at z we know

$$f(z+h) - f(z) - f'(z)\ h = k(h) - f'(z)\ h = \circ(h). \tag{7}$$

From this, it follows that

$$k(h) = f'(z)\ h + \circ(h) = O(h).$$

Thus, we will apply the lemma with the function k in place of F and condition **(iii)** holds.

Rewriting the quantity on the left in (6) as

$$G(k(h)) + (g' \circ f)(z)\ [f(z+h) - f(z) - f'(z)\ h]$$

we can now use the triangle inequality directly:

$$\begin{aligned}
|(g \circ f)(z+h) - (g \circ f)(z) - (g' \circ f)(z)\ f'(z)\ h| &= |G(k(h)) + (g' \circ f)(z)\ [f(z+h) - f(z) - f'(z)\ h]| \\
&\le |G(k(h))| + |(g' \circ f)(z)|\ |f(z+h) - f(z) - f'(z)\ h| \\
&= \circ(h) + |(g' \circ f)(z)|\ |f(z+h) - f(z) - f'(z)\ h| \\
&\qquad\qquad \text{by Lemma 1} \\
&= \circ(h) + |(g' \circ f)(z)|\ \circ(h) \\
&\qquad\qquad \text{by (7)} \\
&= \circ(h). \qquad \square
\end{aligned}$$

There we go: Definitely less than a page.

3 Part (c)

The last part of this problem is about computing chain rules in terms of the standard complex differential operators

$$\frac{\partial}{\partial z} = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)$$

and

$$\frac{\partial}{\partial \bar{z}} = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right).$$

Recall that if $g \circ f$ is differentiable (and we have shown in parts **(a)** and **(b)** above that it is) then we should have

$$\frac{\partial(g \circ f)}{\partial z} = (g \circ f)' \qquad \text{and} \qquad \frac{\partial(g \circ f)}{\partial \bar{z}} = 0.$$

As I have posed the problem, we should be able to show

$$\frac{1}{2}\left(\frac{\partial(g \circ f)}{\partial x} - i\frac{\partial(g \circ f)}{\partial y}\right) = (g \circ f)' \tag{8}$$

and

$$\frac{1}{2}\left(\frac{\partial(g \circ f)}{\partial x} + i\frac{\partial(g \circ f)}{\partial y}\right) = 0. \tag{9}$$

Let's see if we can do this. Let $f = u + iv$ as usual, and set $g = \alpha + i\beta$. Then we have

$$g \circ f = \alpha(u, v) + i\beta(u, v), \tag{10}$$

so by the real chain rule

$$\frac{\partial(g \circ f)}{\partial x} = \alpha_u u_x + \alpha_v v_x + i(\beta_u u_x + \beta_v v_x) \tag{11}$$

and

$$\frac{\partial(g \circ f)}{\partial y} = \alpha_u u_y + \alpha_v v_y + i(\beta_u u_y + \beta_v v_y). \tag{12}$$

Thus,

$$\begin{aligned}\frac{1}{2}\left(\frac{\partial(g \circ f)}{\partial x} - i\frac{\partial(g \circ f)}{\partial y}\right) &= \frac{\alpha_u u_x + \alpha_v v_x + i(\beta_u u_x + \beta_v v_x) - i[\alpha_u u_y + \alpha_v v_y + i(\beta_u u_y + \beta_v v_y)]}{2} \\ &= \frac{\alpha_u u_x + \alpha_v v_x + \beta_u u_y + \beta_v v_y + i(\beta_u u_x + \beta_v v_x - \alpha_u u_y - \alpha_v v_y)}{2}.\end{aligned}$$

Now we have two sets of Cauchy-Riemann equations, one for f and one for g. These tell us

$$\begin{cases} u_x = v_y \\ u_y = -v_x \end{cases} \qquad \text{and} \qquad \begin{cases} \alpha_u = \beta_v \\ \alpha_v = -\beta_u \end{cases}$$

Making these substitutions, we get

$$\begin{aligned}\frac{1}{2}\left(\frac{\partial(g \circ f)}{\partial x} - i\frac{\partial(g \circ f)}{\partial y}\right) &= \frac{\alpha_u u_x - \beta_u v_x - \beta_u v_x + \alpha_u u_x + i(\beta_u u_x + \alpha_u v_x + \alpha_u v_x + \beta_u u_x)}{2} \\ &= \alpha_u u_x - \beta_u v_x + i(\beta_u u_x + \alpha_u v_x).\end{aligned}$$

Now, looking at this, we should ask if (8) holds. In terms of our real and imaginary parts given in (10) and the chain rule proved in part (b), the right side of (8) becomes

$$(g \circ f)' = (\alpha_u + i\beta_u)(u_x + iv_x) = \alpha_u u_x - \beta_u v_x + i(\beta_u u_x + \alpha_u v_x).$$

Thus, we have established (8).

Next, we can compute

$$\frac{1}{2}\left(\frac{\partial(g\circ f)}{\partial x}+i\frac{\partial(g\circ f)}{\partial y}\right)=\frac{\alpha_u u_x+\alpha_v v_x+i(\beta_u u_x+\beta_v v_x)+i[\alpha_u u_y+\alpha_v v_y+i(\beta_u u_y+\beta_v v_y)]}{2}$$
$$=\frac{\alpha_u u_x+\alpha_v v_x-\beta_u u_y-\beta_v v_y+i(\beta_u u_x+\beta_v v_x+\alpha_u u_y+\alpha_v v_y)}{2}.$$

Evidently, the Cauchy-Riemann equations now give

$$\frac{1}{2}\left(\frac{\partial(g\circ f)}{\partial x}-i\frac{\partial(g\circ f)}{\partial y}\right)=\frac{\alpha_u u_x-\beta_u v_x+\beta_u v_x-\alpha_u u_x+i(\beta_u u_x+\alpha_u v_x-\alpha_u v_x-\beta_u u_x)}{2}$$
$$=0,$$

and we see (9) also holds for the composition of holomorphic functions.

Finally, then, let us consider the expressions

$$\frac{\partial g}{\partial z}\frac{\partial f}{\partial z}+\frac{\partial g}{\partial \bar{z}}\frac{\partial \bar{f}}{\partial z}. \qquad \text{and} \qquad \frac{\partial g}{\partial z}\frac{\partial f}{\partial \bar{z}}+\frac{\partial g}{\partial \bar{z}}\frac{\partial \bar{f}}{\partial \bar{z}}.$$

as considered also in the problem. Since all the conjugate derivates vanish for holomorphic functions, the second expresion is zero matching the value in (9). The first expression is

$$\frac{\partial g}{\partial z}\frac{\partial f}{\partial z}.$$

At this point, we can simply recognize that

$$\frac{\partial g}{\partial z}=g'=\alpha_u+i\beta_u \qquad \text{and} \qquad \frac{\partial f}{\partial z}=f'=u_x+iv_x$$

for holomorphic functions, so the first expression matches the value in (8), and we have established

$$\frac{\partial(g\circ f)}{\partial z}=\frac{\partial g}{\partial z}\frac{\partial f}{\partial z}+\frac{\partial g}{\partial \bar{z}}\frac{\partial \bar{f}}{\partial z} \qquad \text{and} \qquad \frac{\partial(g\circ f)}{\partial \bar{z}}=\frac{\partial g}{\partial z}\frac{\partial f}{\partial \bar{z}}+\frac{\partial g}{\partial \bar{z}}\frac{\partial \bar{f}}{\partial \bar{z}} \tag{13}$$

for the composition of holomorphic functions. The problem is complete as I have posed it for Assignment 2.

4 Stein and Shakarchi Exercise 1.8

It may be noticed that Stein and Shakarchi assert something rather more general: They say the formulas (13) hold even when g and f are considered as complex functions (not necessarily complex differentiable) with real and imaginary parts given by continuously differentiable real valued functions. In the notation of (10) Shakarchi and Stein claim we only need to assume $u, v \in C^1(U)$ and $\alpha, \beta \in C^1(V)$ where $V = \{(u, v) \in \mathbb{R}^2 : u + iv \in W\}$ and $U = \{(x, y) \in \mathbb{R}^2 : x + iy \in \Omega\}$ as usual. Let's see if this is true.

We have to be a little careful here because when Stein writes, for example,

$$\frac{\partial g}{\partial z}, \tag{14}$$

he does not mean the expression

$$\frac{1}{2}\left(\frac{\partial(g\circ f)}{\partial x}-i\frac{\partial(g\circ f)}{\partial y}\right)=\frac{\alpha_u u_x+\alpha_v v_x+\beta_u u_y+\beta_v v_y+i(\beta_u u_x+\beta_v v_x-\alpha_u u_y-\alpha_v v_y)}{2}$$

we computed above. What is intended by (14) is the derivative with respect to the complex variable $z = u + iv$ as an argument in g, that is

$$\frac{1}{2}\left(\frac{\partial g}{\partial u} - i\frac{\partial g}{\partial v}\right).$$

We should also recall (or expect) that when an expression like (14) appears in a "chain rule" like one of the expressions in (13), the person who wrote it has in mind the presence of a composition:

$$\frac{\partial g}{\partial z} \circ f = \frac{\partial g}{\partial z}(u + iv).$$

So thinking of $z = x + iy$ in (14) can get you in trouble. One should just think something like: The partial derivative of a function with respect to z is "partial with respect to real part variable minus i times partial with respect to imaginary part variable." With this in mind, we can compute:

$$\frac{\partial g}{\partial z} = \frac{1}{2}[\alpha_u + i\beta_u - i(\alpha_v + i\beta_v)] = \frac{1}{2}[\alpha_u + \beta_v + i(\beta_u - \alpha_v)].$$

Similarly,

$$\frac{\partial f}{\partial z} = \frac{1}{2}[u_x + v_y + i(v_x - u_y)].$$

Therefore,

$$\begin{aligned}\frac{\partial g}{\partial z}\frac{\partial f}{\partial z} = \frac{1}{4}\,[&\alpha_u u_x + \alpha_u v_y + \beta_v u_x + \beta_v v_y - (\beta_u v_x - \beta_u u_y - \alpha_v v_x + \alpha_v u_y)\\ &+i(\alpha_u v_x - \alpha_u u_y + \beta_v v_x - \beta_v u_y + \beta_u u_x + \beta_u v_y - \alpha_v u_x - \alpha_v v_y)]\\ = \frac{1}{4}\,[&\alpha_u u_x + \alpha_u v_y + \beta_v u_x + \beta_v v_y - \beta_u v_x + \beta_u u_y + \alpha_v v_x - \alpha_v u_y\\ &+i(\alpha_u v_x - \alpha_u u_y + \beta_v v_x - \beta_v u_y + \beta_u u_x + \beta_u v_y - \alpha_v u_x - \alpha_v v_y)]\,.\end{aligned}$$

For the first formula in (13) we also want to compute

$$\frac{\partial g}{\partial \bar{z}}\frac{\partial \bar{f}}{\partial z}.$$

This is a computation much like the one we've just done:

$$\frac{\partial g}{\partial \bar{z}} = \frac{1}{2}[\alpha_u + i\beta_u + i(\alpha_v + i\beta_v)] = \frac{1}{2}[\alpha_u - \beta_v + i(\beta_u + \alpha_v)].$$

Similarly,

$$\frac{\partial \bar{f}}{\partial z} = \frac{1}{2}[u_x - iv_x - i(u_y - iv_y)] = \frac{1}{2}[u_x - v_y - i(v_x + u_y)].$$

It follows that

$$\begin{aligned}\frac{\partial g}{\partial \bar{z}}\frac{\partial \bar{f}}{\partial z} = \frac{1}{4}\,[&\alpha_u u_x - \alpha_u v_y - \beta_v u_x + \beta_v v_y + (\beta_u v_x + \beta_u u_y + \alpha_v v_x + \alpha_v u_y)\\ &+i(-\alpha_u v_x - \alpha_u u_y + \beta_v v_x + \beta_v u_y + \beta_u u_x - \beta_u v_y + \alpha_v u_x - \alpha_v v_y)]\\ = \frac{1}{4}\,[&\alpha_u u_x - \alpha_u v_y - \beta_v u_x + \beta_v v_y + \beta_u v_x + \beta_u u_y + \alpha_v v_x + \alpha_v u_y\\ &-i(\alpha_u v_x + \alpha_u u_y - \beta_v v_x - \beta_v u_y - \beta_u u_x + \beta_u v_y - \alpha_v u_x + \alpha_v v_y)]\,.\end{aligned}$$

Summing up we obtain

$$\frac{\partial g}{\partial z}\frac{\partial f}{\partial z} + \frac{\partial g}{\partial \bar{z}}\frac{\partial \bar{f}}{\partial z} = \frac{1}{2}\,[\alpha_u u_x + \beta_v v_y + \beta_u u_y + \alpha_v v_x - i(\alpha_u u_y - \beta_v v_x - \beta_u u_x + \alpha_v v_y]\,. \tag{15}$$

Now my computation above for Problem 7 of Assignment 2 as I posed it, up until the Cauchy-Riemann equations were used, gives

$$\frac{\partial(g \circ f)}{\partial z} = \frac{\alpha_u u_x + \alpha_v v_x + \beta_u u_y + \beta_v v_y + i(\beta_u u_x + \beta_v v_x - \alpha_u u_y - \alpha_v v_y)}{2}.$$

Comparing this to the expression in (15), we see that we have established the first chain rule of (13).

The computation for my problem above also gives

$$\frac{\partial(g \circ f)}{\partial \bar{z}} = \frac{\alpha_u u_x + \alpha_v v_x - \beta_u u_y - \beta_v v_y + i(\beta_u u_x + \beta_v v_x + \alpha_u u_y + \alpha_v v_y)}{2}.$$

A similar computation yields the second chain rule of (13). Perhaps I leave it to you.