

Misconceptions And Difficulties With The Chain Rule

Tangül Uygur and Aynur Özdaş
tuygur@anadolu.edu.tr , aozdas@anadolu.edu.tr

INTRODUCTION

Derivative is an inherently difficult concept for many students. Especially whenever the function considered is a composite function, the difficulties students often encounter increase and get worse. Tall (1993, p.19) indicated “The Leibniz notation $\frac{dy}{dx}$ proves to be almost

indispensable in the calculus. Yet it causes serious conceptual problems. Is it a fraction, or a single indivisible symbol?” According to Tall, one difficulty with the notion of the Chain

Rule is the dilemma of whether the du can be cancelled in the equation $\frac{dy}{dx} = \frac{dy}{du} \times \frac{du}{dx}$. Let

consider the case in which a second order derivative of a composite several variable function is required. Most students often find themselves within a notational and conceptual complexity in this case. As cited in Artigue (1991, p.170), Poincaré wrote about the second order derivative:

As soon as one passes to derivatives of the second order, one swims in absurdity: let z be a function of a variable y , which is itself a function of x ; I write:

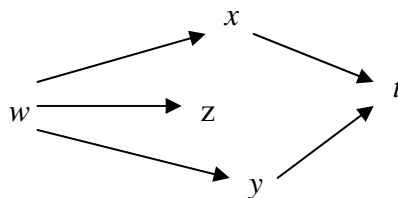
$$\frac{d^2z}{dx^2} = \frac{d^2z}{dy^2} \times \frac{dy^2}{dx^2} + \frac{dz}{dy} \times \frac{d^2y}{dx^2}$$

In this formula I write d^2z twice, and the symbol has two different meanings...

Clark, Cordero, Cottrill, Czarnocha, Devries, John, Tolias, and Vidakovic (1997) studied understanding (genetic decomposition) of the Chain Rule and its applications. They came to the conclusion that the difficulties with the chain rule for a large number of students could be attributed to student difficulties in dealing with the composition and decomposition of functions. Cottrill (1999) studied also the correlation between a student’s understanding of the composition of functions and understanding the chain rule. In his study, there was a small amount of evidence supporting the hypothesis which states that the understanding of composition of functions is key to understanding the chain rule. Cottrill indicated a new study was needed to address this hypothesis.

II.METHOD

The purpose of this study is to determine the misconceptions and difficulties the students taking Analysis IV course in Elementary Mathematics Education Program at Anadolu University encounter with the chain rule. Note that, while the chain rule was taught to these students, an arrow diagram was used for applying this rule. Let us give an example of the arrow diagram. Let $w = f(x, y, z)$, $x = g(t)$, $y = h(t)$. The arrow diagram for these equations is given below:



When having to apply the chain rule, one can use the arrow diagram by following the algorithm that: identifying the different branches of the arrow diagram going from the dependent variable to the selected variable; then for each branch, writing the derivative or partial derivative associated to each arrow and multiplying them; finally adding the expressions obtained for each branch.

This study is part of a rather extensive investigation, aimed to assess the participants’ facility with not only the chain rule but also the related notions as Leibniz notation, composite function,

partial differentiation and the arrow diagram which the researchers uses in teaching the chain rule. The study was administered during the second semester of the 2004-2005 academic year.

A test including four essay type items was prepared for purposive sampling. This test was administered to 27 students who had taken Analysis III course. The items of the test were as follows:

1. Let $w = \cos(u + v)$, $u = \tan x$ and $v = \ln x$. Find the derivative of w respect to x .
2. If $y=f(u)$ and $u=g(x)$, evaluate second order derivative of y respect to x .
3. Given $w = \sin(\tan x + e^x)$, find derivative of w respect to x .
4. If $w=f(x,y)$ and $x=x(t)$, evaluate second order derivative of w respect to t .

To analyze the result of this test, following skills coded as seen below were considered; then 27 students were grouped according to them.

[Drv]: Evaluating derivative of special functions (such as in the first and third items)

[Chain 1]: Applying the chain rule for the first order derivative.

[Product]: Applying the product rule to the derivative function.

[Chain 2]: Applying the chain rule to derivative functions.

One student volunteered to participate was randomly selected from each group. Distributions of the students to the groups according to the skills in which they have weaknesses are shown below with an asterisk.

Number of the students	[Drv]	[Chain 1]	[Product]	[Chain 2]	Selected Participant
4	*			*	P1
7					P2
6				*	P3
4		*	*	*	P4
6			*	*	P5

Table 1

After analysing the results of the test and selecting the participants, the first draft of the interview form was prepared. Then this draft was piloted with a student by one of the researchers. Interview tasks and their answers were checked for errors and validity by the researchers. After some revisions, interview tasks were constructed as follows:

- Find the first and second order derivatives of y respect to x for the below cases.
 1. $y = x^4 + 3x + 1$
 2. $y = (3x + 2)^2$
 3. $y = (g(x))^3$
 4. $y = f(u), u = g(x)$
- Find the first and second order derivatives of z respect to x for the below cases.
 5. $z = \sin(x^3 + 2y^2 + xy)$
 6. $z = f(u), u = g(x, y)$
 7. $z = u^2 + v^2, u = x + 2y, v = \cos(x + y)$
 8. Let $z=f(u,v)$, $u=g(x,y)$, and $v=h(x,y)$, Evaluate the derivative of z respect to x .
 9. If $w=f(u,v,z)$, $v=g(u,x)$, and $x=h(u)$, evaluate the derivative of w respect to u .

Semi-structured interviews taking one to two hours were conducted with each participant selected from 27 students by the same researcher. An interview calendar was prepared with the participants of the study in order to use their and researchers' time economically. Prior to conducting the interviews, each time the researcher explained the purpose of the interview, where the interview data were to be used, the time needed, and the confidentiality of their names. The interviews were audio and video taped with the camera. Questions were written on a worksheet by the researcher while she was asking them. It was required from participants to write the answers on this worksheet besides expressing them verbally and not to erase writings even if they were wrong. The audio tapes were transcribed to written form and read carefully several times by regarding the worksheets. We studied student responses to the first to sixth tasks. Seventh, eight, and the ninth tasks were prepared to investigate the effects of the arrow diagram the researchers used in teaching the several variable

composite functions' derivative in Analysis III course in previous semester. As expected, the participants who gave responses to these tasks had used this diagram. Therefore, responses to these tasks were not considered in this study. Transcriptions were coded for key events by the researchers and worksheets were also used as written data especially for the codes related to the notations. Moreover, the video tapes and worksheets were logged to ease reference to key events found in the audio transcripts. Some sub-questions related to the notions: (I) the chain rule, (II) composite function, (III) relationship between the chain rule and the composition and (IV) the Leibniz notation were asked at the convenient moment according to the way the interview was going. Coding was done according to the following questions, related to misconceptions and difficulties emerged during the interviews.

1. Did the student call the memorized rules for finding derivatives of special composite functions (such as the power functions, the trigonometric functions, etc.) differently from "the chain rule"?
2. Could the student use the chain rule as a second way for finding derivatives including some special cases?
3. Did the student have difficulties with evaluating the second order derivative of the special composite function by using the Leibniz notation in spite of the fact that she could evaluate this second order derivative by the thought "the outside function is then composed with the inside function"?
4. Did the student use the Leibniz notation by modifying in his/her own way?
5. Did the student call and explain the notion of the composition in his/her own way while she was finding the derivative of some special composite functions?
6. Did the student avoid using the Leibniz notation for finding the second order derivative, although she had general formula of the chain rule with this notation?
7. Did the student simplify the general formula of the chain rule?
8. Was the student successful in evaluating the second order derivative of an abstract composite function?

First, second, third, fifth, and the seventh questions were coded with **Yes** or **No**. The fourth and sixth questions were coded with **No Use (NU)** if the student was not able to use this notation for second order derivative, otherwise coded with **Yes** or **No**. Eighth question was coded with **Unsuccessful**, **Prompted**, or **Successful**. These codes were used in a manner similar to the codes of Cottrill(1999). The codes are summarized in Table 2.

	1	2	3	4	5	6	7	8
P1	Y	Y	Y	Y	Y	Y	Y	P
P2	N	N	N	N	N	Y	Y	P
P3	N	N	Y	N	Y	Y	N	P
P4	Y	Y	Y	N	Y	NU	N	U
P5	Y	Y	Y	NU	Y	NU	N	U

Table 2

III.RESULTS

As seen the codes for the eighth question, two of the five participants could not evaluate the second order derivative of the abstract function whereas three of the five were able to evaluate with prompt. If the codes of the third question are considered, it can be realized that except P2, all participants had difficulties with evaluating the second order derivative of the special composite function by using the Leibniz notation in spite of the fact that they could evaluate this second order derivative by the thought "the outside function is then composed with the inside function". For instance the part of the interview with P1, which relates this issue, is as follows:
 (The researchers explanations are given in parenthesis and writings of the participants found necessary to give are given in curly parenthesis below.)

I: Okay. Let's find the second order derivative of y respect to x now. ($y = (3x + 2)^2$)

P1: dy is divided by dx^2 , we have already found the first order derivative therefore the derivative of this will be 18. But it would seem that we can find in this way (She is indicating the general formula of the chain rule)

I: How can you find in this way?

P1: I will compare also this way (applying the chain rule) to this (writing the composite function as $y = f(x)$ then calculating the derivative), dy divided by du , again in there du divided by, ..., I will

write the second order of this in order to again get this $\frac{dy}{dx^2}$ when we simplify $\left\{ \frac{dy}{du} \cdot \frac{du}{dx^2} \right\}$. The

derivative of this (u^2) was $2u$, the second order derivative of this ($3x$) will be 0 because I should write the derivative of 3, when I find the second order derivative, this will be 0.

I: All right...you had found the second order derivative in another way, what did you find?

P1:in this way I found only dy dived by du , then I realized that this (her second way) was wrong, it is true in this way (her first way) now

I: Okay. What is this 18?

P1: 18 is the derivative of this, that is, the derivative of the first order derivative.

I: In other words?

P1: It is becoming the second order derivative which is the coefficient of $3x$.

I: Which is true according to you?

P1: **This is definitely true (indicating 18)**

I: Why? How can you be sure?

P1: Because hmmm..., **when we applied the rules**, the first order derivative was true..., I know this is certainly true in this case and when **I find the derivative of this again in the frames of rules, I will find the coefficient.**

Then P1 calculated the second order derivative by writing as $\left(\frac{dy}{dx} \right)' = 18$.

As realized codes for the first question, three of the five participant applied the memorized rules as different from the chain rule and called as "changing variable", "convert the simple form", "to make simpler". What P5 said about this was that:

I: Okay, you multiplied these, what is the name of this rule?

P5: Is it the chain rule?

I: Yes, all right. **When you wrote $z' = \cos u \cdot u'$, did you apply the chain rule?**

P5: **We didn't use the chain rule, ..., no didn't use.**

Even if four of the five participants had the notion of composition, they were not able to transfer this knowledge to apply the chain rule and they called the notion of the composition as one function inside another function or in his/her own way. P5, one of them, expressed this as follows:

I: Okay, when you apply the chain rule?

P5: The chain rule, the function respect to another variable which the function don't depend on, ...or how can I say..., if we get the function another form, for example in there, I join z to u and the derivative respect to x is required, therefore we have to pass from u in order to reach x , I apply the chain rule in this way.

If the codes for sixth question is considered, it can be seen that all participants could not use or avoided to use the Leibniz notation for second order derivative, although they had general statement (the general formula) of the of the chain rule with this notation. For instance, for $y=f(u)$, $u=g(x)$, P3 wrote the second order derivative as $y'' = (f(u))' \cdot (g(x)')^2 + (g(x))' \cdot f(u)'$

,although she wrote the first order derivative as $\frac{dy}{dx} = \frac{dy}{du} \cdot \frac{du}{dx}$. P2, who had difficulty only in

notation category and needed to prompt about evaluating the second order of abstract function also

wrote the solution of this question similar to P3: $\frac{dy^2}{d^2x} = f''(u).g'(x).g'(x) + g''(x).f'(u)$. P2 was also one of two participants, simplifying the general formula of the chain rule:

I: Okay. Let $y=f(u)$, $u=g(x)$. Now find the derivative of y respect to x .

P2: Hmm...(thinking)we are applying the chain rule (laughing)

I: All right, apply now. What was the chain rule?

P2: Hmm..., they cancelled each other, for example $\frac{dy}{dx}$ we were writing such that $\frac{dy}{du}$ multiply $\frac{du}{dx}$, these du 's cancelled each other therefore $\frac{dy}{dx}$ remains. This makes it simpler.

IV.DISCUSSION

After completion of data analysis, we came to the conclusion that almost all students are able to evaluate derivative of special composite function by memorized rules, but most of them calculate these derivatives without the conscious use of the chain rule. Moreover, they could provide the general statement of the chain rule and write down the formula, but a few students could explain the connection between the general statement of the chain rule and memorized rules. Because they could not relate the memorized rules to the chain rule, they call and explain the composite functions in their own way while they are applying the memorized rules to the special composite functions. Another aspect of these misconceptions is notation. Some students perceive the Leibniz notation as a fraction and they simplify the formula. Almost all students avoid using the Leibniz notation. Whenever a student who has these misconceptions and difficulties attempt to evaluate the second derivative of an abstract composite function, he finds himself in a notational complexity. Furthermore, the student

who could not perceive the notation of $\frac{d}{dx}$ as an operator and the notation like $\frac{dy}{dx}$ as a function is

not able to evaluate the second order derivative of the special composite function by using the Leibniz notation in spite of the fact that he could evaluate this second order derivative by the thought "the outside function is then composed with the inside function". If a student has knowledge of the derivative function, he could succeed the second order derivative with the prompt about the Leibniz notation. As seen from Table 2, all participants had difficulty with the second order derivative. Except two of them, who could not use the Leibniz notation, remaining three of them could succeed the second order derivative of the abstract composite function with prompt. Consequently, we thought that the role of the notations in the success with the chain rule is as great as the role of the notions of composite and derivative functions. We suggest that in teaching the concept of the chain rule, more emphasis should be given to using the Leibniz notation meaningfully, to relating the special cases to the general statement of the chain rule and the abstract cases in order not to meet such misconceptions. Another important point in teaching this rule is prompting the students with relating the composite and derivative function notions to the various especially abstract problem situations which embody the chain rule concept.