Students' arguments about the growth of a two-variable function

Andreas Bergwall

Örebro University, Department of Science and Technology, Sweden; andreas.bergwall@oru.se

Calculus is a central part of the curriculum for tertiary educations in mathematics, science, and technology. At its core lies the concept of derivative, which is known to be problematic for many students. As the corresponding multi-variable concepts of partial derivative, gradient, and directional derivative are not mathematically equivalent, it is essential for students to learn their relations and what they represent geometrically. In this paper, 20 students' written solutions to an exam problem about the growth of a two-variable function are studied. The warrants they present for their claims are characterized in terms of which representations, concepts, connections, and calculations they use. The findings indicate that students who solve the problem by calculation of directional derivatives are less explicit with their warrants than students who rely on properties of the gradient vector. While the first group only uses algebraic representations, the second combines algebraic and graphical representations.

Keywords: Tertiary mathematics education, gradient, directional derivative, Toulmin model.

Introduction, review and aim of the study

Millions of students around the world study calculus at university. In educations in science, technology, and economics calculus are often mandatory. In parts of the world, such as Europe, calculus is also a dominant part in upper secondary mathematics. On this level the approach is usually informal with focus on applications and procedural skills (Törner et al., 2014). The transition to the more formal mathematical analysis at university level is often problematic for students as it involves a shift from a pragmatic to a deductive praxeology (Job & Schneider, 2014). While there is a rather extensive body of educational research with focus on basic problems of teaching and learning calculus, at upper secondary and introductory tertiary level, there has been a shortage of studies that go beyond the early topics of calculus (Rasmussen et al., 2014). Multivariable calculus is such a topic.

The existing research shows that the concepts of multivariable analysis are abstract and difficult for students (Martínez-Planell et al., 2015a). On the one hand research has highlighted the importance of being able to use multiple representations, on the other that most students' have difficulties with graphical representations of two-variable functions (Kabael, 2011). While textbooks tend to assume that concepts such as slope naturally extend to multivariable settings, this has been showed to be a source of learning difficulties (McGee & Moore-Russo, 2015).

One of the most central concepts of calculus is that of derivative, which in the single-variable setting represents instantaneous growth and slope of tangent lines. When generalizing to the multivariable case several concepts arise, which are not mathematically equivalent to each other, such as partial derivative, gradient, and directional derivative. Connected to them are properties such as partial differentiability, differentiability, directional differentiability, and C^1-regularity. Understanding and using the relations between these objects and properties, and what they represent geometrically, are an essential part of learning multivariable differential calculus. However, little is known about how students understand basic concepts such as partial derivatives, differentials, directional derivatives, and gradients. It is documented that they have difficulties interpreting the geometric meaning of partial derivatives and differentials (Martínez-Planell et al., 2015a), and problems understanding that the gradient of a two-variable function is a two-dimensional vector, and not a normal to the function's graph in three-dimensional space (Cline et al., 2012). Even students who have successfully finished a course in multivariable calculus often lack a deeper understanding of directional derivatives and rely on memorized facts and formulas (Martínez-Planell et al., 2015b). Regarding the logical relations between continuity, partial differentiability, differentiability, and directional differentiability, the study by Lankeit and Biehler (2019) indicates that students often can decide which implications that are true, but that they have difficulties producing correct arguments for their claims. With the present study we aim to contribute on how learners of multivariable calculus understand and use the various multivariable concepts of derivative in a simple situation. More precisely, in relation to a specific exam problem, we aim to answer the following research question: What characterizes the warrants students offer for their claims about growth of a two-variable function?

Method and analytic approach

Analytic framing

This study is based on a qualitative analysis of students' written solutions to a task given as part of a written exam. As we are interested in students' argumentation, their solutions will be analyzed using the three fundamental parts of Toulmin's model of argument (Toulmin, 2003): *ground*, *warrant*, and *claim*. In this model, the claim (or conclusion) refers to an assertion of some kind, and the ground (or data) is the facts the assertion is based on. The warrant is what links ground and claim. In relation to the student solutions analyzed in the present paper, the information given in the task is considered the ground. The answer the student produce is the claim. Anything additional the student presents to back up the answer is the warrant. In this respect, the complete solution is the unit of analysis. As mathematical calculations and reasoning usually proceed stepwise, we will allow ourselves to also make this analysis on a more fine-grained level: the outcome of one step of the solution will be considered both a claim of that step and ground for the next step, and so on. This means that a solution usually includes a sequence of grounds, warrants, and claims.

As research has shown the importance of using multiple representations, this is one of the aspects of students' solutions that will be analyzed. We will distinguish between *verbal*, *algebraic*, and *graphic* representations. Verbal representations refer to texts written in ordinary language, algebraic representations refer to numeric and algebraic calculations and derivations, and graphic representations refer to geometric figures and diagrams.

We will also focus on content specific aspects in students' solutions. Their warrants and claims will be described in terms of which mathematical concepts they make use of (such as partial derivatives, gradients, and directional derivatives). How the students connect these (and other) concepts to each other will be inferred from how they combine them in their solutions. What mathematical calculations and operations they use will also be accounted for (e.g., partial differentiation, dot multiplication of vectors, normalization of vectors, comparison of magnitudes and angles, substitution of variables with numbers).

Context and data sample

The data sample consists of 20 students' written solutions to the following task:

> Assume that grad $f(x,y) = (-3x^2y + 2xy, -x^3 + x^2 - 2y)$. Viewed from the point $(1,-1)$, in which of the directions $(1,1)$ and $(2,1)$ does $f(x,y)$ increase most rapidly?

The task was one of a set of elementary tasks at a written re-exam on a course in multi-variable calculus, given in the third semester of a civil engineering program in computer science at a Swedish university. The exam also included a set of more advanced and complex tasks. The general instructions included that the students should justify their reasoning, include all essential steps of calculations, draw clear figures, and present exact answers. The assessment criteria included that solutions should be easy to follow, be based on a working solution strategy, be presented using suitable forms of representations, include arguments for the essential steps, and not lead to unreasonable or absurd answers.

During earlier semesters the students had had courses in discrete mathematics, linear algebra, and single-variable calculus, as well as courses in programming and physics. The multi-variable calculus course included differential calculus, integral calculus, and vector calculus. The differential calculus part included limits, partial derivatives, gradients, directional derivatives, the chain rule, Taylors formula, and local and global optimization, all with focus on the two- and three-dimensional cases. Properties such as continuity, partial differentiability, differentiability, and C^1-regularity, and how they are related, were included in the course but the students were not required to learn the proofs. Differentials were not given explicit attention.

In the course (textbook and lectures) the gradient was defined in cartesian coordinates as grad $f(x,y) = (f'_x(x,y), f'_y(x,y))$. Directional derivatives were defined for unit vectors $\boldsymbol{v}$:

$$f'_{\boldsymbol{v}}(a,b) = \lim_{h \to 0} \frac{f\big((a,b) + h\boldsymbol{v}\big) - f(a,b)}{h}$$

The students were presented the proof (using the chain rule) for the identity $f'_{\boldsymbol{v}}(a,b) = \text{grad } f(a,b) \cdot \boldsymbol{v}$ (which holds if f is differentiable at (a,b)) and its corollary (which follows by using the Cauchy-Schwarz inequality) that grad $f(a,b)$ is the direction of most rapid growth (steepest accent) at (a,b).

The lectures included calculations of growth in specific directions and determination of direction of most rapid growth by use of algebraic methods, but never comparisons of growth in different directions. Despite that, the course teacher (the author of this paper) considered the exam task a routine task and expected most students to solve it by comparing the values of the directional derivative $f'_{\boldsymbol{v}}(1,-1) = \text{grad } f(1,-1) \cdot \boldsymbol{v} = (1,2) \cdot \boldsymbol{v}$ with respect to the two normalized vectors $\boldsymbol{v} = \frac{1}{\sqrt{2}}(1,1)$ and $\boldsymbol{v} = \frac{1}{\sqrt{5}}(2,1)$.

Analytic procedures

The students' written solutions were first sorted according to which solution strategy they applied. Nine of the 20 solutions used a working strategy in the sense that the applied algorithms and procedures would surely lead to a correct answer (to any task of this kind) if all computational details were done correctly. However, these solutions often included minor computational or notational errors or lacked important steps. To characterize warrants used to back up claims, the nine solutions were analysed in terms of what representations, concepts, connections, and calculations that were presented to justify the answer to the task. Through this analysis two subcategories of solutions were formed. Examples and a detailed analysis are presented in the result section.

The remaining eleven solutions (all based on non-working strategies) were also analysed according to use of representations, concepts, connections, and calculations. This led to the realization that eight of them would have worked had the formulation of the task (i.e. the ground) been slightly different. We believe these solutions can bring some insight in typical misunderstandings or misconceptions about growth of two-variable functions. Examples are given in the result section.

Analysis and results

As we are not aiming for general conclusions about students' argumentation, we will not provide a full account of all data but focus on detailed accounts of three selected solutions, which represent the three most common solution strategies and illustrate the variation in the data. The first two strategies lead to correct solutions (if calculations are carried out correctly), while the third does not. We will also mention some details of the other incorrect solutions in the data. The student solutions have been translated to English and reproduced in the handwriting of the author of this paper. Solution steps are numbered according to the disposition in the authentic student solutions.

Two working strategies

Four student solutions included the following steps: calculation of $\text{grad}(1,-1)$ by substituting (x,y) for $(1,-1)$ in the expression for grad $f(x,y)$, calculation of the directional derivatives $f_v'(1,-1)$ for the vectors (1,1) and (2,1) by using the formula $f_v'(a,b) = \text{grad } f(a,b) \cdot \boldsymbol{v}$, and selection of the direction that yielded the largest directional derivative. An example is given in Figure 1.

1. grad $f(1,-1) = (1,2)$
2. $f_v'(1,1) = \text{grad } f(1,-1) \cdot \frac{1}{\sqrt{2}}(1,1) = \frac{1}{\sqrt{2}}(1,2)\cdot(1,1) = \frac{3}{\sqrt{2}}$
3. $f_v'(2,1) = \text{grad } f(2,1) \cdot \frac{1}{\sqrt{5}}(2,1) = \frac{1}{\sqrt{5}}(1,2)\cdot(2,1) = \frac{4}{\sqrt{5}}$
4. Normalized direction vektors (1,1) and (2,1) and got $\frac{1}{\sqrt{2}}(1,1)$ and $\frac{1}{\sqrt{5}}(2,1)$ respectively.
5. Seen from the point (1,-1) it is the direction (1,1) f(x,y) grows most rapidly in.

Figure 1: Correct solution based on comparison of directional derivatives

The warrant that students who use this strategy presents for their choice of direction vector is that the directional derivative is largest in that direction, that is, their claim rests on comparison of directional derivatives. Breaking down the argument, it has two parts. First, the students make claims about the values of the directional derivatives. The warrants for these claims are algebraic and consist of calculations of dot products (Figure 1, steps 2 and 3). Second, these values are used as grounds for choosing a direction vector (Figure 1, step 5). In the solution in Figure 1, no warrant is presented for this second part. In one of the other solutions in this category, the student underlined the largest value but made no further comment. One student backed up this final step by calculating and comparing $\left(3/\sqrt{2}\right)^2$ and $\left(4/\sqrt{5}\right)^2$.

None of the students who produced solutions in this category explicitly pointed out that directional derivatives correspond to growth in a specific direction, or that it is directional derivatives they are calculating. However, the use of the notation $f_v'(a,b)$ and/or the use of the formula grad $f(a,b) \cdot \boldsymbol{v}$ (as in Figure 1, steps 2 and 3), makes it plausible to assume that this is what they do. Either way, the students obviously see no need to make this an explicit part of their warrant.

Summarizing, students, who make claims about direction of most rapid growth based on calculation of directional derivatives, provide explicit warrants for the values of the directional derivatives in the form of algebraic computations. They use the algebraic connection between gradient and directional derivative expressed in the formula $f_v'(a,b) = \text{grad}\, f(a,b) \cdot \boldsymbol{v}$. Mistakes observed in the solutions (aside from minor arithmetic errors) are that direction vectors are not normalized, errors in the calculation of the dot product, and notational errors.

Five students chose another approach and provided solutions that included the following steps: calculation of $\text{grad}(1,-1)$ by substituting (x,y) for $(1,-1)$ in the expression for grad $f(x,y)$, marking of the gradient vector $\text{grad}(1,-1)$ and the direction vectors $(1,1)$ and $(2,1)$ in a diagram, and selection of the direction that deviates the least from $\text{grad}(1,-1)$. An example of this kind of solution is given in Figure 2.

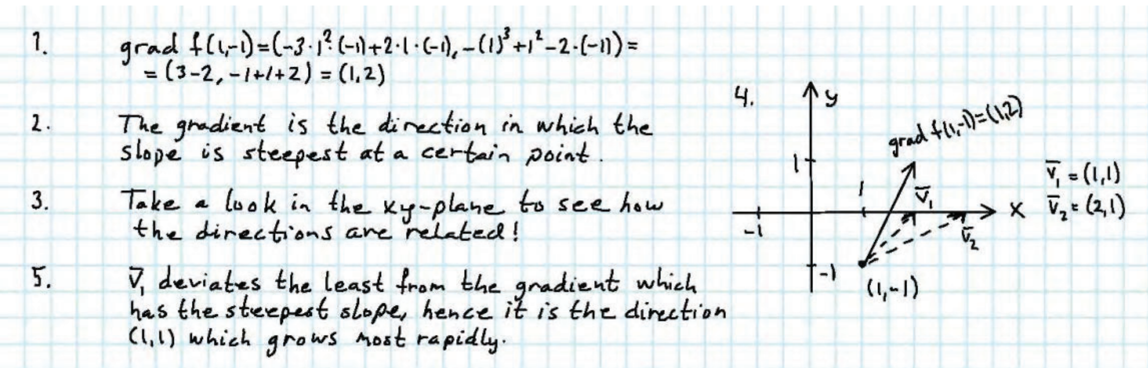


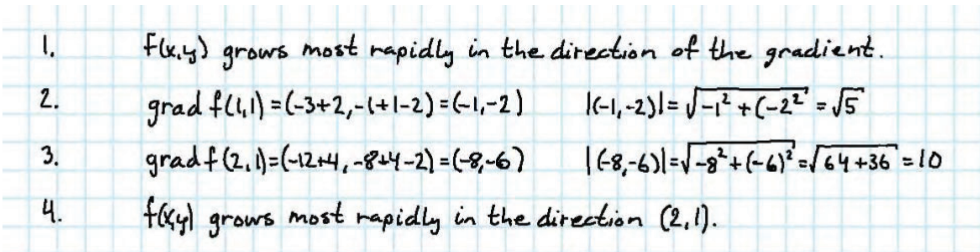
Figure 2: Correct solution based on comparison with the gradient vector

The warrant that students who use this strategy presents for their choice of direction vector is that the chosen direction deviates the least from the direction of most rapid growth, that is, their claim rests on comparison of directions. Breaking down the argument, it has two parts. First, the students make a claim about which of all possible directions that gives the most rapid growth. The warrant for this is the algebraic computation of $\text{grad}(1,-1)$ (Figure 2, step 1) combined with a verbal statement that the gradient provides the direction of most rapid growth (Figure 2, steps 2). This statement is not backed up with a refence to a theorem or a derivation. Second, the students choose the direction vector they claim to deviate the least from the gradient vector. This claim is expressed verbally (Figure 2, step 5), and a warrant is given graphically as a diagram in which the gradient and the direction vectors are marked from the same point (Figure 2, step 4). Not all students explicitly tell that they pick the vector which deviates the least from the gradient, and only one student explicitly says that it is the smallest *angular* deviation that is essential. Unfortunately, students will answer correctly even if they compare the norms of the vector differences. Thus, it is not possible to know if they have compared angles or vector differences. None of the students provide a warrant for why the least deviation from the gradient gives the most rapid growth.

Summarizing, students, who make claims about direction of most rapid growth based on comparison with the direction of the gradient vector, provide explicit warrants for the gradient as the direction of most rapid growth in the form of verbal statements, and explicit warrants for which direction that deviates the least from the gradient vector in the graphical form of a diagram. They use the connections between the gradient vector and the direction of most rapid growth, and the assumption that the lesser a direction deviates from the gradient, the more rapid is the growth. An interesting feature of this solution strategy is that the students are able to answer the question about growth without actually calculating the growth. This simplifies the task from a computational perspective and may explain why none of the students made any computational errors.

Non-working strategies

The most common non-working strategy to solve the task included the following steps: calculation of the gradient vector at the *points* $(1,1)$ and $(2,1)$, calculation of the length of $\mathrm{grad}(1,1)$ and $\mathrm{grad}(2,1)$, and selection of the vector/point that yielded the largest length. Five solutions were of this kind. An example is given in Figure 3.



1. $f(x,y)$ grows most rapidly in the direction of the gradient.
2. $\mathrm{grad}\, f(1,1) = (-3+2, -1+1-2) = (-1,-2)$ $\quad |(-1,-2)| = \sqrt{-1^2 + (-2^2)} = \sqrt{5}$
3. $\mathrm{grad}\, f(2,1) = (-12+4, -8+4-2) = (-8,-6)$ $\quad |(-8,-6)| = \sqrt{-8^2 + (-6)^2} = \sqrt{64+36} = 10$
4. $f(x,y)$ grows most rapidly in the direction $(2,1)$.

Figure 3: Incorrect solution based on comparison of the lengths of gradient vectors

This strategy would be correct if the task had asked at which of the points (1,1) and (2,1) the function has its largest maximum rate of change. Without additional data on the student reasoning (such as interview data) it is impossible to judge whether the students misunderstood the formulation of the task or if they do not understand the difference between rate of change at different points and rate of change in different directions. It could also be that students lack fundamental knowledge about the difference between points and vectors in the plane. Thus, it is difficult to say if the flaw lies in the ground or the warrant of their arguments. Anyway, the warrant that students who use this strategy presents for their choice of direction/point is that the gradient has its largest length at that point, that is, their claim rests on comparison of the lengths of gradient vectors. In the solution above the student's claim about most rapid growth in the gradient direction is given without warrant (Figure 3, step 1). The gradient vectors' coordinates and lengths are claims warranted by algebraic computations (Figure 3, steps 2 and 3). The vector lengths can be assumed to be the ground for the final claim (Figure 3, step 4), even though there is no explicit warrant offered for this claim.

Summarizing, students, who make claims about direction of most rapid growth based on comparison of lengths of gradient vectors at different points, provide explicit warrants for the gradient as the direction of most rapid growth in the form of verbal statements, and explicit warrants for the lengths of gradient vectors in the form of algebraic computations. They use (at least implicitly) the connections between the gradient vector and the direction of most rapid growth, and the length of the gradient vector and the rate of change.

The remaining six student solutions were incorrect and differed considerably from each other. In several of them the students mixed up the meanings of points and vectors in some way. In one of them the student argued as if the phrase "the directions (1,1) and (2,1)" in the formulation of the task should be interpreted as "the directions towards the points (1,1) and (2,1)". Aside from this, that particular solution was very similar to the one exemplified in Figure 2.

Concluding remarks

With this study we contribute with findings on students' argumentation about how the growth of a two-variable function relate to its gradient, a field sparsely studied in the research literature. Based on an analysis of 20 written solutions to an exam problem, two common working solution strategies, and one non-working, have been identified. These strategies have been described with focus on the warrants students offer for their claims. While existing literature tends to focus on students' difficulties and shortcomings, this study also describes correct student solutions. Regarding student difficulties, the findings confirm existing research.

It is impossible to infer whether students who chose one of the working strategies also were familiar with the other one. To find out more about this, a follow-up study with classroom observations and interviews is planned. But if we assume that students chose the strategy they are most confident with, there are some possible differences between the two categories of students that are worth discussing. The students who chose to compare directional derivatives showed knowledge about how to use the gradient to compute them. They also seemed to rely on procedures, algebraic representations, and memorized formulas. If this is a correct interpretation it is in line with the findings reported by Martínez-Planell et al. (2015b). The warrants they offered for their claims were only in form of algebraic computations. The second group, who compared directions with the gradient vector, may not have been confident with the concept of directional derivative, but they included a multitude of forms of representations in their arguments, and showed a deeper conceptual understanding of the core concepts. In addition to algebraic computations, they offered warrants in the form of diagrams, and they verbally stated important facts about the connection between the gradient and the direction of most rapid growth. Contrary to the findings in Cline et al. (2012) they did not mistake the gradient for being a three-dimensional vector. In a sense it was the students who did *not* show any knowledge about directional derivatives that showed the deepest understanding of how growth of a two-variable function is related to its gradient, and it was those students who could argue for their claims using a multitude of representations. This indicates the importance, and difficulty, of designing learning activities that promote learning of algebraically efficient methods (such as the formula for directional derivatives) as well as geometrical understanding and skills in algebraic and non-algebraic reasoning and communication.

Eleven out of the 20 analyzed solutions did not apply a working strategy, even though the task is an elementary multivariable calculus task. Even among those who chose a correct approach minor errors of various kinds were common. This confirms the findings of Martínez-Planell et al. (2015a) that fundamental concepts of multivariable calculus are abstract and difficult for students, even after they have attended a multivariable calculus course.

The fact that a directional derivative depends on a function, a point, and a vector, contributes to its complexity. In this study, only the vector varied, albeit some students interpreted it as if the point varied. This suggests that design research, using principles from variation theory, can offer a way forward to develop (theories for) teaching and learning about directional derivatives.

References

Cline, K., Parker, M., Zullo, H., & Stewart, A. (2012). Addressing common student errors with classroom voting in multivariable calculus. *Primus*, *23*(1), 60–75. https://doi.org/10.1080/10511970.2012.697098

Job, P., & Schneider, M. (2014). Empirical positivism, an epistemological obstacle in the learning of calculus. *ZDM - The International Journal of Mathematics Education, 46*(4), 635–646. https://doi.org/10.1007/s11858-014-0604-0

Kabael, T. U. (2011). Generalizing single variable functions to two-variable functions, function machine and APOS. *Educational Sciences: Theory and Practice, 11*(1), 484–499.

Lankeit, E., & Biehler, R. (2019). *Students' work with a task about logical relations between various concepts of multidimensional differentiability.* Paper presented at the Eleventh Congress of the European Society for Research in Mathematics Education.

Martínez-Planell, R., Gaisman, M. T., & McGee, D. (2015a). On students' understanding of the differential calculus of functions of two variables. *The Journal of Mathematical Behavior, 38*, 57–86. https://doi.org/10.1016/j.jmathb.2015.03.003

Martínez-Planell, R., Gaisman, M. T., & McGee, D. (2015b). Student understanding of directional derivatives of functions of two variables. In T. G. Bartell, K. N. Bieda, R. T. Putnam, K. Bradfield, & H. Dominguez (Eds.). *Proceedings of the 37th annual meeting of the North American chapter of the international group for the psychology of mathematics education* (pp. 355-362). Michigan State University.

McGee, D. L., & Moore-Russo, D. (2015). Impact of explicit presentation of slopes in three dimensions on students' understanding of derivatives in multivariable calculus. *International Journal of Science and Mathematics Education, 13*(2), 357–384. https://doi.org/10.1007/s10763-014-9542-0

Rasmussen, C., Marrongelle, K., & Borba, M. C. (2014). Research on calculus: What do we know and where do we need to go? *ZDM - The International Journal of Mathematics Education, 46*(4), 507–515. https://doi.org/10.1007/s11858-014-0615-x

Toulmin, S. E. (2003). *The uses of argument*. Cambridge University Press.

Törner, G., Potari, D., & Zachariades, T. (2014). Calculus in European classrooms: Curriculum and teaching in different educational and cultural contexts. *ZDM - The International Journal of Mathematics Education, 46*(4), 549–560. https://doi.org/10.1007/s11858-014-0612-0