Communications

Is mathematical productivity a quasi-OCD?

KYRIAKOS PETAKOS

American psychologist G. Mandler poses the question "Is it useful to think of math anxiety as a quasi-neurosis?" (1989, p. 241). Mandler's question inspired me to reflect on mathematical productivity or mathematical creativity as sort of a mental disorder, specifically Obsessive Compulsive Disorder (OCD).

The DSM-5 [1] defines Obsessive-Compulsive Disorder, in part, as:

> A. Presence of obsessions, compulsions, or both:
>
> Obsessions are defined by (1) and (2):
>
> 1. Recurrent and persistent thoughts, urges, or impulses that are experienced, at some time during the disturbance, as intrusive and unwanted, and that in most individuals cause marked anxiety or distress.
>
> 2. The individual attempts to ignore or suppress such thoughts, urges, or images, or to neutralize them with some other thought or action (i.e., by performing a compulsion).
>
> Compulsions are defined by (1) and (2):
>
> 1. Repetitive behaviors (e.g., hand washing, ordering, checking) or mental acts (e.g., praying, counting, repeating words silently) that the individual feels driven to perform in response to an obsession or according to rules that must be applied rigidly.
>
> 2. The behaviors or mental acts are aimed at preventing or reducing anxiety or distress, or preventing some dreaded event or situation; however, these behaviors or mental acts are not connected in a realistic way with what they are designed to neutralize or prevent, or are clearly excessive. (§300.3)

Mathematical productivity, which for me includes the ability to conceive and prove new theorems, has many parallels with the concept of obsession as defined in DSM-5. You begin by using tenets and principles that have been stated and proven previously, and from which you are expected to reach a new conclusion. The passion that characterizes your approach fuels your desire to obtain a new result, one you have already shaped in your mind. But every human being's thinking has its limits. At some point the thoughts become so strong that they penetrate your whole system and, at first imperceptibly but later consciously, continue in other parts of your social life. The hitherto positive thoughts of discovery dominate other parts of your life; they become intrusive, unwanted thoughts. Moments of your life, supposed moments of happiness, are marred by thoughts of mathematics.

Does that not resemble the mechanism that gives birth to the unwanted, recurrent thoughts that form the obsessive component in OCD? Parallel to Mandler's question "To what extent is such a reaction similar to the bland accounts of neurotic patients who deny or avoid their real-life difficulties?" (p. 241) we can ask "To what extent are the thoughts when one seeks to create new mathematics similar to the intrusive ones of OCD patients who avoid doing or continuously do certain actions induced by this set of thoughts?"

The connection between the OCD's compulsion and mathematics is less clear, but still present. You may be uncertain of the reliability of your results. To prevent or reduce anxiety you consult a colleague to acquire a feeling of certainty of what you have accomplished so far. Sharing the glory of the discovery or the disappointment of its failure is at the same time a method of sharing the uncertainty you have experienced during your research progress. Without a person with whom you can discuss your problem and thinking, you might seek reassurance from yourself, your own logic and knowledge, a cognitive process intertwined with mixed emotions. It is a common mathematical practice to go over and check every one of your steps. That entails going over and over again certain operations, proofs you have already performed. When you feel comfortable about something, a positive emotion dominates you, the joy of creation, the joy of discovery. But once again your state of joy appears to be evanescent. You have already become accustomed to checking and checking, to repeatedly questioning the validity of your work. Is that not a typical compulsive behavior?

While engaged in our mathematical research, the mixture of emotions is a reality that cannot be ignored. We love or hate ourselves according to whether we have achieved or not our goal. We feel joy or we feel sadness. Ultimately, the desire to discover something new is fueled by our past achievements, and the wonder with which we regard our findings, at least at the moments when uncertainty has not yet overwhelmed us.

I see here a possible association of OCD with not only mathematical creativity among researchers, but also in teaching. A good teacher should be able to understand a wide range of perspectives, maintaining at the same time her loyalty to the main goal of her mission, which is a relational understanding of mathematics (Skemp, 1976). As the teacher delves into her own knowledge and mastering of mathematics to find ways to guide her students to a relational understanding of mathematics, she cannot really avoid being obsessed about it. At the same time she faces many obstacles that might cause her to doubt the value of her efforts, including questioning of her goal by commentators outside the school, and pressure to show immediate results.

The teacher faces further dangers, related to the compulsive component of OCD. While the researcher obtains a kind of closure with publication of a result, and can move on to new problems, the teacher often teaches the same topic to different classes in parallel or in different semesters. This offers many opportunities to question her approach, to repeat her teaching with small variations over and over. This might improve the quality of her teaching, but be detrimental to herself. The parallel with OCD offers a new lens through which to examine with tension in teaching.

If, as I claim, mathematical creativity can be seen as a sort of Obsessive Compulsive Disorder, could the experiences of mathematicians inform the treatments of psychologists, and could the work of psychologists assist the productive work of mathematicians? Could methods employed by mathematicians to control their anxiety when encountering obstacles in their research be used by psychologists to ameliorate analogous situations occasioning symptoms of OCD?

I wrote this communication as an exploration of an idea, based on parallels I saw between mathematical activity and feature of OCD. By chance, as I was writing it I read of a clear, if extreme, example of the parallel I was imagining. According to a BBC online article (Keating, 2019), in 2002 Jason Padgett was mugged by two men, resulting in a brain injury that brought on OCD. Unable to leave his home, and finding his perceptions of the world around him fundamentally altered, he developed an interest in mathematics and physics, subjects he had dismissed as useless before he was attacked. Through the internet he developed this interest and took to trying to capture in drawings the way he now saw the world around him. "I had literally a thousand or more drawings of circles, fractals, every shape that I could manage to draw. It was the only way I could manage to communicate effectively what I was seeing." A chance meeting with a physicist led Padgett to enroll in a mathematics course, and through mathematics he became able to venture out into the world.

This story is unusual, but shows how the interrelatedness between the OCD and the mathematical productivity might be usefully explored by both mathematicians and psychologists. Mathematics might improve the lives of people who suffer from mental disorders, and psychology might enhance mathematical research.

Acknowledgement

I am grateful to Alf Coles for providing me with resources and encouragement to submit this communication.

Note

[1] American Psychiatric Association (2013) *Diagnostic and Statistical Manual of Mental Disorders.* Washington, DC: American Psychiatric Association Publishing.

References

Keating, S. (2019) The violent attack that turned a man into a math genius. Online at http://www.bbc.com/future/story/20190411-the-violent-attack-that-turned-a-man-into-a-maths-genius

Mandler, G. (1989) Affect and learning: reflections and prospects. In McLeod, D.B. & Adams, V.M. (Eds.) *Affect and Mathematical Problem Solving,* pp. 237–244. New York: Springer.

Skemp, R.R. (1976) Relational understanding and instrumental understanding. *Mathematics Teaching* **77**(1), 20–26.

Characterising mathematical activities promoted by Fermi problems

LLUÍS ALBARRACÍN, JONAS B. ÄRLEBÄCK

Some questions that arise in everyday situations, such as 'How long will it take to get to the airport?' can be solved by making a quick estimate. In other cases, we can ask ourselves about situations that we have never considered before, where we are interested in obtaining a first rough answer. This would be the case if we ask how many planes are flying all over the world at a specific time are, or the amount of CO_2 emissions we could avoid in a city if gasoline-powered cars were replaced by electric ones. These two questions are examples of Fermi problems. They owe their name to Enrico Fermi, who used this particular type of problems both in his scientific work and as a university teacher. Fermi problems, being smaller, more well-defined and delimited contextualised problems and not real-world problems in all their complexity, have been considered to be miniature-modeling problems (Robinson, 2008).

Fermi problems can be solved by decomposing the original problem into simpler sub-problems, solving the original question through reasonable estimates. This way of solving realistic problems has been used in mathematics education to introduce mathematical modelling to primary education students (Peter-Koop, 2009), to use real-world knowledge to support mathematical learning with secondary education students (Albarracín & Gorgorió, 2014; Ärlebäck, 2009) and at university level to assess modelling processes, such as validating models (Czocher, 2018).

Although educational research on Fermi problems generally emphasises estimation, it has been suggested that the activity of estimation can be replaced by other (classroom) activities to find the numerical information needed to solve the problem (Sriraman & Knott, 2009). In this communication we identify different types of mathematical activities that can be derived from the use of Fermi problems in the existing research, with the goal to develop a characterisation framework that can be used both for designing research interventions and as an analytical tool in research.

Fermi problem activity templates

In this section we present a characterisation tool called the *Fermi problem Activity Template* (FpAT) that focuses on the structure of the Fermi problem. It can be used to describe either (a) the result of an a priori analysis of the Fermi problem split into interconnected subproblems needed to be solved; or (b) a structure that describes the activities students actually engage in when solving a Fermi problem. Hence, FpAT can be used both in the design of research interventions and as an analytical tool.

Part of our inspiration for how we have come to represent the FpAT categorisation comes from Anderson and Sherman (2010). They proposed a representation based on an a priori analysis of the problem at hand that structures the solving

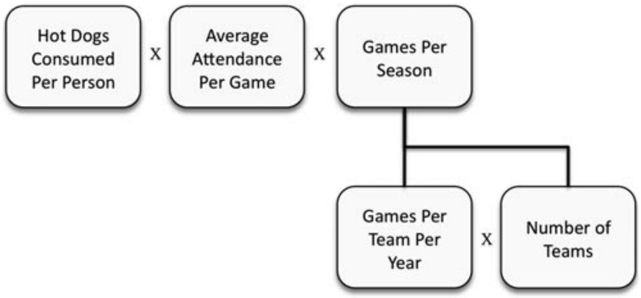


Figure 1. The MBL hotdog Fermi problem as presented in Anderson and Sherman (2010, p. 37).

process by explicitly differentiating the sub-problems the students have to engage in to solve the problem. The example they discuss in great detail is about estimating the number of hotdogs consumed at the Major League Baseball (MLB) games each season in the US. Differentiating between the values needed to be estimated (such as the number of hotdogs consumed per person and per game) from given values that can be looked up (such as the number of games in a season), Anderson and Sherman (2010) presented the breakdown of the problem as in Figure 1.

One can consider engaging in different types if activities to achieve the unknown numerical values needed in solving the different sub-problems. For example, rather than using a standard way of estimating based on previous experiences, consulting the MLB website might provide the number of games each team plays per year, and to find the number of spectators at a typical game one can turn to newspaper reports. In the case of finding the number of hotdogs consumed per person and game, a survey that generates sufficient data to determine a reasonable estimate can be used. The decision and reason for using one activity rather than another can depend on the solver's understanding of the precision required in the solution of the problem, or be a conscious choice made in the design of the activity in order for the students to work toward given certain mathematical contents and curricular goals.

Mathematical activities promoted by Fermi problems

Some studies have used the so-called MAD framework (Ärlebäck, 2009) to analyse the activities students engage in when solving Fermi problems. This framework characterises the process of solving a Fermi problem in terms of the following sub-modelling activities: *reading, making a model, estimating, calculating, validating* and *writing*. Estimating is the activity originally seen as key to a Fermi problem, but we concur with the suggestion of Sriraman and Knott (2009) that it can be generalised. In the following, we argue and illustrate that it can incorporate other ways for students to get the necessary numerical information they need to answer the problem or related sub-problems. The four types of mathematical activities that we propose are *guesstimation, experimentation, data search* and *statistical data collection.*

To visualise the structure of the problem and the different types of mathematical activities that can be used in solving

Table 1. The components of the Fermi problem Activity Template framework (FpAT).

Activity / representation	**Students obtain the quantity by engaging in …**
 Guess-timation	… a mental process giving a rough solution through guessing and making comparisons based on previous experiences and intuition.
 Experi-mentation	… in-and-out-of-school experimentations and investigations, including making measurements.
 Looking for data	… searching for numerical information in external sources.
 Statistical data collection	… suitable ways of selecting, collecting and analysing statistical data.

the problem, our FpAT framework uses different graphical representations for the four (sub-)activities (see Table 1). The purpose of using different geometrical shapes for the different activities is to enhance readability.

Figure 2 shows the FpAT framework applied to the structure of the Fermi problem presented in Figure 1, illustrating that a statistical survey will be used to find the average number of hotdogs consumed per person in a game, that how many people attend a game will be determined using an estimate, and that the rest of the needed numerical values will be sought in appropriate sources. The different sub-problems identified in the problem are delimited using square brackets in order to clearly convey the structure of the problem.

Using FpAT as an analytic tool

In Albarracín and Gorgorió (2013) one of the problems studied asked individual students for an estimate of the total number of SMS messages sent in a day in Catalonia. Secondary students (12-16 years old) were asked to provide a detailed and concrete plan for estimating the quantities asked for rather than solving the problem. The following excerpt presents one of the proposed solutions given by a

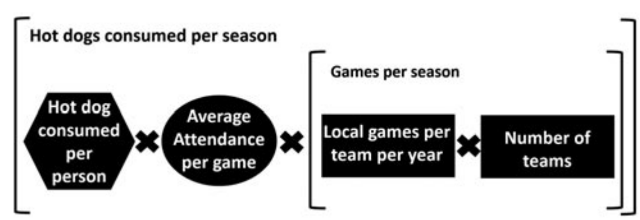


Figure 2. Structure of the MBL Hotdog problem characterised using FpAT.

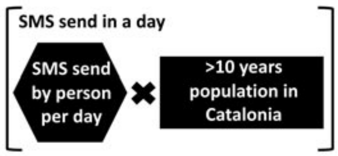


Figure 3. *Structure of student A1's proposed solution to the SMS problem represented using FpAT.*

student (A1). Figure 3 shows the student's solution represented using FpAT.

> I would make a survey of the people I know about how many messages they send per day, trying to take a balanced number of women and men, as well as young people, adults and the elderly. Then, I would make an average and then multiply this by the number of inhabitants of Catalonia between the age ranges that might have a mobile phone.

Student A1's proposal is to use a survey to find out the number of SMS messages sent by the inhabitants of Catalonia in a day. The student considers that the participants surveyed must comply with some restrictions in order for the sample to be adequate, but does not provide any details on how to realise this. Then, when the average number of SMS messages sent in a day per person is known, the student proposes to multiply this number by the number of mobile users in Catalonia.

Figure 4 represents student A2's proposed solution:

> First, it [the survey] will take 10 teenagers, 10 adults, 10 senior citizens, from here I would ask them how many SMS they send per day to each group and I would find the average. Then I would calculate the total average respecting the percentage of people of each type and multiply it by the total number of people in Catalonia.

Student A2 also suggests using a survey to collect data as the basis for his solution, but in contrast to Student A1's proposal suggests stratifying the sample according to his experiences and knowledge about the different age groups' usage of SMS messages. As illustrated by the FpAT in Figure 4, however, Student A2's proposal does not necessarily

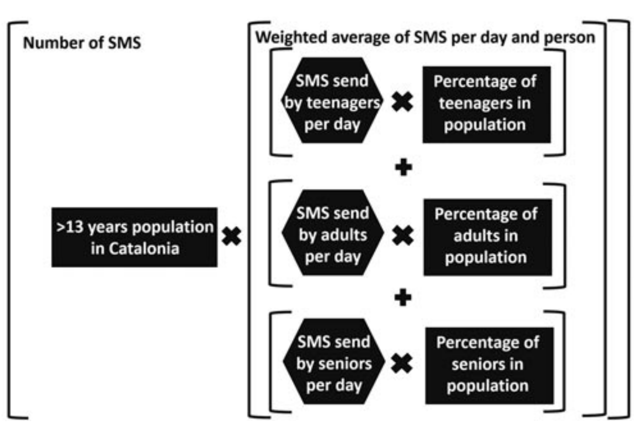


Figure 4. *Structure of student A2's proposed solution to the SMS problem represented using FpAT.*

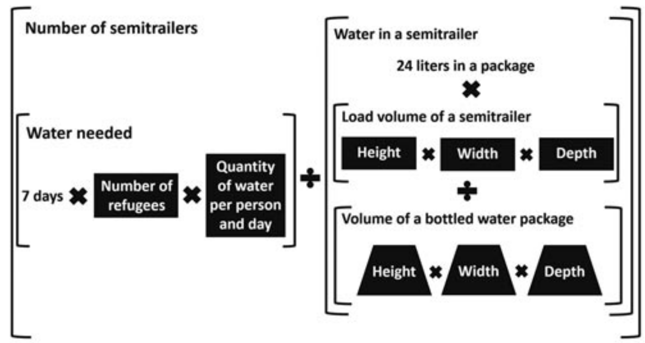


Figure 5. *Structure of problem discussed in Taggart et al. (2007).*

use this stratification in each population group, but rather suggests using the collected data to derive a weighted average. These two examples illustrate that FpAT as an analytical tool of students' work facilitates a straightforward and compact representation of students' (proposed) solutions, as well as making subtle differentiating aspects of the processes (potentially) involved in solving the problem visible, such as in the case of Student A2's suggestion of using a weighted average.

Using FpAT as a tool for task design

To illustrate the potential of the FpAT framework in designing research interventions, we apply the framework to an a priori analysis of an activity proposed in Taggart *et al.* (2007). It asks students "How many semitrailer loads of b ottled water [16.9 fluid ounces] were needed by New Orleans' refugees in the week following Hurricane Katrina?" (p. 166). Taggart *et al.* discuss this question and list a number of sub-questions that need to be answered in order to answer the original question.

Figure 5 presents the FpAT based on the sub-questions in Taggart *et al.,* and we can see that the problem is structured around determining two main quantities: the total amount of water needed and the amount of water that a semi-trailer can transport. However, these two quantities must be determined from a succession of estimates and calculations using more specific quantities. This a priori analysis highlights the different types of activities students can productively engage in when tackling the problem. Hence, given as a modelling problem in a research intervention, and depending on the focus of the research, it can support the research design by preparing and supporting the research setting.

Final thoughts

Fermi problems have a long tradition in certain educational contexts. However, despite various recommendations in the mathematics education literature, this type of problem has not yet to any large degree been implemented in classrooms, nor has its full potential been exploited or researched. As we have argued here, Fermi problems in a modelling context promote multiple types of mathematical activities allowing for a broad range of solutions, classroom discussions and learning at all educational levels, in multiple subjects. Indeed, given that the contexts of Fermi problems often come from the STEM disciplines, Fermi problems have the potential to function as integrating activities, allowing for the connection and transfer of strategies and applications of the different activities from other disciplines into mathematical classrooms (and vice versa). In addition, the different types of activities in the FpAT framework are accessible entrance points to introduce and work with various types of technologies that can be used to obtain or process data, allowing Fermi problems to also be a facilitator of students' ICT literacy. With this in mind, we are hopeful that the FpAT framework can be used as a research tool, both for designing interventions and analysis, especially focusing on what demands, roles, functions and effects different activities have on the problem-solving process, the solutions, and students' learning.

Acknowledgements

Lluís Albarrasín is Serra-Húnter Fellow in Universitat Autònoma de Barcelona. Research supported by the projects EDU2017-82427-R (Ministerio de Economía, Industria y Competitividad, Spain), 2017 SGR 497 (AGAUR, Generalitat de Catalunya) and José Castillejo program (CAS17/00289 reference).

References

Albarracín, L. & Gorgorió, N. (2013) Problemas de estimación de grandes cantidades: modelización e influencia del contexto. *Revista Latinoamericana de Investigación en Matemática Educativa* **16**(3), 289–315.

Albarracín, L. & Gorgorió, N. (2014) Devising a plan to solve Fermi problems involving large numbers. *Educational Studies in Mathematics* **86**(1), 79–96.

Anderson, P. & Sherman, C. (2010) Applying the Fermi estimation technique to business problems. *Journal of Applied Business and Economics* **10**(5), 33–42.

Ärlebäck, J.B. (2009) On the use of realistic Fermi problems for introducing mathematical modelling in school. *The Mathematics Enthusiast* **6**(3), 331–364.

Czocher, J.A. (2018) How does validating activity contribute to the modeling process? *Educational Studies in Mathematics* **99**(2), 137–159.

Peter-Koop, A. (2009) Teaching and understanding mathematical modelling through Fermi-Problems. In Clarke, B., Grevholm, B. & Millman, R. (Eds.) *Tasks in Primary Mathematics Teacher Education,* pp. 131–146. New York: Springer.

Robinson, A.W. (2008) Don't just stand there—teach Fermi problems! *Physics Education* **43**(1), 83–87.

Sriraman, B. & Knott, L. (2009) The mathematics of estimation: possibilities for interdisciplinary pedagogy and social consciousness. *Interchange* **40**(2), 205–223.

Taggart, G.L., Adams, P.E., Eltze, E., Heinrichs, J., Hohman, J. & Hickman, K. (2007) Fermi questions. *Mathematics Teaching in the Middle School* **13**(3), 164–167.

Signed numbers and signed letters in algebra

AMIR ASGHARI

The Glossary of the Common Core State Standards for Mathematics (2010) defines an integer by *a way that it can be expressed.*

> **Integer.** A number expressible in the form *a* or –*a* for some whole number *a*.

This definition hints at three mathematical objects, a signed number, a signed letter and an integer, that it is useful to distinguish. A *signed number* is a specific natural number prefixed by the sign – or + [1]. A *signed letter* is a letter prefixed by the sign – or + to denote an unspecified negative or positive quantity. A letter may stand for a parameter or a variable, and I will use the expressions 'signed parameter' and 'signed variable' to reflect the role of the letter in the situation under discussion. In both signed numbers and signed letters, the + sign is often omitted (as it is in the Common Core definition), and here my focus is on the – sign. Signed numbers and signed letters have two parts, the number or letter that indicates a specified or unspecified number greater than 0, and the sign. An integer, on the other hand, is a unity, a number that, when added to another integer, gives the sum 0.

Thus for example, '−3' as a signed number is the number 3 prefixed by the sign – . But, as an integer, '−3' is the solution of ? + 3 = 0; "all we need to know about −3 is that when you add 3 to it you get 0" (Gowers, p. 26). Similarly, '–*a*' as a signed letter is the natural number a prefixed by the sign –. But, as an integer, '$-a$' is the solution of ? + a = 0, and *a* itself might be less than 0.

The purpose of this short communication is to provide some historical evidence to show the role of signed letters in the development of algebra. In particular, some ingenious and correct ways that mathematicians used signed letters for expressing generality are shown, that now seem cumbersome and inefficient. It is suggested that we should be more aware and explicit about the use of signed letters and numbers in schools today.

Negative numbers at the advent of algebra

The historical rejection of integers as numbers is well known. Eventually, their utility was one of the most important reasons for their acceptance and development. For example, as Hefendehl-Hebeker (1991) writes "using negative numbers one can quickly and effectively solve problems that make sense only for positive numbers and have only positive solutions" (p. 32). An important factor in the gradual acceptance of negative numbers was due to the usefulness of the rules of signs applied to signed numbers and signed letters. According to Nathaniel Hammond's 1742 *The Elements of Algebra in a New and Easy Method* [2] the rules are as follows :

> When the quantities to be multiplied have like signs, that is, they are both affirmative, or both negative, then set or join the letters together, and to them prefix the sign +, which will be the product required. (Art. 9, p. 18)
>
> When the signs of the two quantities that are to be multiplied are one affirmative and the other negative, then multiply the quantities as before directed, but to their product prefix the negative sign, or – . (Art. 16, p. 26)

With a slight change of language (*e.g.,* using 'positive' instead of 'affirmative'), this is "the rule for multiplying" given by the most popular dedicated educational site in the UK [3]:

> When the signs are different the answer is negative.
>
> When the signs are the same the answer is positive.

Signed numbers and signed letters with the correct use of the rules of signs suffice for getting the correct answer in calculations and working with algebraic expressions. They even work well for finding the solutions of equations, but they are inefficient for expressing the equations generally for which we need to use parameters.

Signed parameters

> We have got this equation $x^2 + 3x - 10 = 0$, whence $x^2 + 3x = 10$ (Hammond, Art. 70, p. 249).

For us, $x^2 + 3x - 10 = 0$ is a special case of the general form of quadratic equations $x^2 + bx + c = 0$, with $b = 3$ and $c = -10$. But for Hammond, it is a special case of $x^2 + bx = c$ with $b = 3$ and $c = 10$. The other two forms [4] of quadratic equations are $x^2 - bx = c$ and $x^2 - bx = -c$.

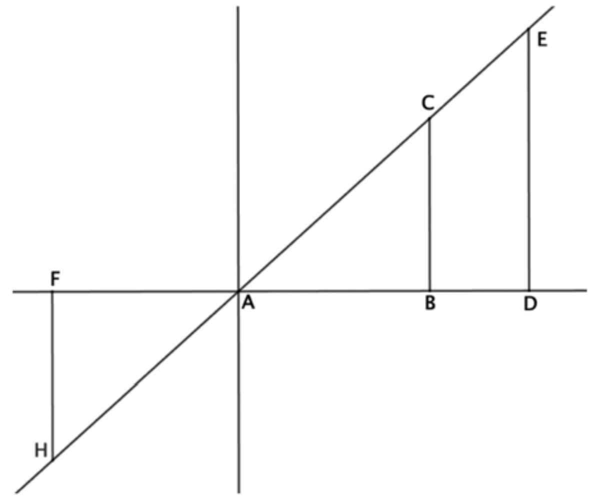
With *signed parameters* (*i.e.,* signed letters to represent the coefficients), not only the general expressions of the quadratic equations but also the relevant theorems are divided into signed cases. For example, instead of the general theorem that the sum of the roots of $x^2 + bx + c = 0$ is $-b$, there are two theorems :

> If the co-efficient of [x] has the sign, + the sum of both the roots will be the same as the co-efficient, but will have the sign –.
>
> And, if the co-efficient of [x] has the sign –, [...] then the sum of both the roots or values, will be the same as the co-efficient, but will have the sign +.
>
> Therefore having found any one root, the other is easily found. (Art. 62, p. 188)

Let us consider the form where the co-efficient of x has the sign +. To find the other value of the unknown quantity, the first value is added to the co-efficient of x, "and to their sum prefix the sign –"(Art. 63, p. 188). For example, having found $x = 5$ as one of the roots of $x^2 + 10x = 75$, we add 5 to 10. The sum is 15. "Prefix to this 15 the sign – and this is the other value of x, that is $x = -15$." (Art. 63, pp. 188–189).

It would be constructive to compare Hammond's approaches with ours. For us, the sum of the roots of $x^2 + bx + c = 0$, is $-b$. Let the roots be x_1 and x_2, we can write $x_1 + x_2 = -b$. Now it does not matter if the root we have is positive or negative. Since x_1 admits both positive and negative numbers, we substitute the found root into $x_1 + x_2 = -b$ for x_1 and knowing b (that also could be positive or negative), we find x_2. This line of thought, that may seem very natural to the reader, is beyond Hammond's *Elements of Algebra.* Hammond's letters cannot admit varying negative quantities. They can represent *fixed unknown quantities* in equations. But, if –15, for example, is a solution to an equation, it must be "an *imaginary* value of *x,* it being absurd for a positive quantity to be equal to a negative one." (Art. 63, p. 189). Though it is "absurd" for the letter representing the root to be equal to a negative quantity, it is in line with the other rules of algebra, and hence, its acceptance is somehow justified: "we shall find this *imaginary* value of *x,* if we proceed by division according to the directions at Art. 61." (Art. 63, p. 189).

The whole process of finding the roots is based on the established rules of signs applying to signed numbers and signed letters. Though "absurd", the fixed value may turn out to be negative, but it is never *chosen* to be negative in the sense that a variable may admit negative numbers as well as positive. So for example, "No quantities connected by the sign + only [...], can be equal to nothing. That is, it cannot be [...] $a + b = 0$ (Hammond, Art. 77, p. 326).



Through the points A, C, draw an indefinite right line HE; this will be the locus of the two equations $y = \frac{ax}{b}$ and $y = -\frac{ax}{b}$. For, taking any line $AD = x$, and drawing DE parallel to BC, it will be $DE = \frac{ax}{b} = y$. And taking $AF = -x$, and drawing FH parallel to BC, it will be $FH = -\frac{ax}{b} = y$.

Figure 1. Agnesi's description of the graph of y = (a/b)x (redrawn and simplified, from Book I, Section III, Art 117; Colson, p. 93).

Before reading the next section, please imagine yourself in the Hammond's time and try to draw the graph of the equation $y = -x$, that is $x + y = 0$.

Signed variables

To observe a true use of letters as variables we have to move a few years forward from Hammond to 1748, the publication year of Maria Gaetana Agnesi's *Instituzioni Analitiche* (translated as *Analytical Institutions for the Use of Italian Youth* by John Colson, 1801 [5]). It includes some of the earliest examples of using co-varying quantities for expressing algebraic equations of curves, and "is significant for its clarity of exposition and its widespread influence as a textbook" (Boyer, 2012, p. 177). In it we can also find an explicit introduction to signed numbers and letters in algebra:

> Positive and negative quantities in algebra are distinguished by means of certain marks, or signs, which are prefixed to them. To positive quantities the sign +, or *plus,* is prefixed; to negative quantities the sign –, or *minus.* (Agnesi, 1748, Book I, Section I, Art. 3; p. 2 in Colson)

With $+a$ only denoting a positive quantity and $-a$ only a negative quantity, the general equation of a straight line (that is, for us $y = mx + b$), should be divided into different types, that according to Agnesi are as follows [6]: $y = mx$; $y = -mx$; $y = mx + c$; $y = -mx - c$; $y = mx - c$; $y = -mx + c$ (Book I, Section III, Art. 117; pp. 92-93 in Colson).

This is another example of signed parameters used for expressing a generality in *the absence of a letter admitting both positive and negative numbers.* But then in the equation of each of the lines above there are also two letters x and y that are not for representing fixed numbers. They are *variable quantities,* "because the value of one of the unknown quantities may be varied an infinite number of ways, so, in like manner, the value of the other may be as often varied" (Book I, Section III, Art. 111; p. 90 in Colson). The locus of each of the equation above is a right 'line', but not the one we are used to. The equation $y = mx$ is a ray in the first quadrant and $y = -mx$ is a ray in the third quadrant. In other words, to algebraically describe the line HE in Figure 1, we need two equations rather than one.

This is not, as Boyer suggests, "some of the old errors with respect to negative coordinates" (Boyer, 2012, p. 178). Rather, it is a *correct* way of dealing with the situation when the variable quantities are *signed variables.* Nowadays, it would be an error if our students assumed that $y = mx$ only represents positive y's and $y = -mx$ only negative y's. We expect them to see that the point (x, y) may lie in any of the quadrants, both x and y are unsigned variables admitting both positive and negative numbers, m is a parameter representing both positive and negative numbers, and hence, $y = mx$ is the general equation of lines passing through the origin.

School mathematics

We now present graphs of equations on the Cartesian plane with positive and negative numbers on the axes, and expect that our students will not experience difficulties accepting unsigned variables. After all, even a line as simple as $y = x$, assumes both positive and negative numbers for the variables involved. However, Christou and Vosniadou's (2005) study shows that students have a "strong tendency to interpret literal symbols as standing only for natural numbers" (p. 456). Even those few students who assigned a negative value to a literal symbol, say a, did it by "by putting a minus sign in front of it" (p. 455), changing it to $-a$.

Historically, the acceptance of unsigned variables was not easy. And it seems that even now it is not easy for our students. This difficulty may become less surprising if we consider that it is quite possible to do arithmetic, solve equations, and work with algebraic expressions, solely using signed numbers and signed letters (for this purpose, Hammond's book is still usable). And many students only encounter positive variables (mainly variables that admit natural numbers, *e.g.,* in so-called matchstick patterns) for a long time, before experiencing variables in general.

We should not assume that the transition to unsigned variables happens spontaneously for our students. The fact that variables can take negative values should be addressed directly. It may help students to see how unsigned variables simplify things, as they can use a *single* equation $y = mx + b,$ for a line in the plane, while Agnes needed six.

Notes

[1] Signed numbers can also represent negative and positive rational or irrational numbers, but for simplicity I will at first consider signed numbers corresponding to integers.

[2] Available online at https://archive.org/details/elementsalgebra02hamm-goog. I have updated Hammond's notation slightly (*e.g.,* using x as the variable, and writing x^2 instead of xx) and changed the capitalisation of words to follow current standards.

[3] BBC Bitesize at https://www.bbc.co.uk/bitesize/guides/z77xsbk/revision/3

[4] There should be the fourth form $x^2 + bx = -c$. But Hammond does not discuss it. In cases where one of the roots is negative he justifies excluding it, so I believe that he was aware of the fourth case but chose to overlook it as it has only negative "absurd" roots.

[5] Available online at https://archive.org/details/analyticalinsti00masegoog

[6] Agnesi represented m by the fraction a/b and hence used c for the y-intercept.

References

Boyer, C.B. (2012) *History of Analytic Geometry.* Mineola, NY: Dover.

Common Core State Standards Initiative. (2010) *Common Core State Standards for mathematics.* Online at http://www.corestandards.org/Math/Content/mathematics-glossary/glossary/

Christou, K.P. & Vosnaidou, S. (2005) How students interpret literal symbols in algebra. In Bara, B.G., Barsalou, L. & Bucciarelli, M. (Eds.) *Proceedings of the 27th Annual Conference of the Cognitive Science Society,* pp. 453–458. Mahwah, NJ: Lawrence Erlbaum Associates.

Gowers, T. (2002) *Mathematics: A very short Introduction,* Vol. 66. Oxford: Oxford Paperbacks.

Hefendehl-Hebeker, L. (1991) Negative numbers: obstacles in their evolution from intuitive to intellectual constructs. *For the Learning of Mathematics* **11**(1), 26–32.