Part 3: Mathematical Thinking and Rigor, First Steps

Luis Moreno-Armella

Corey Brady

Part 3: Mathematical Thinking and Rigor, First Steps

Luis Moreno-Armella*; Corey Brady

CINVESTAV-IPN; Southern Methodist University

Abstract: Agustín thus sought to abandon the *unsafe* use of inductive methods in algebra, as Leonardo had employed them, and to rather *recover* Euclid's ideal that is crystallized in the *Elements of Geometry*. Nevertheless, his idea of a limit is grounded in the dynamism of motion or a process: It is an *embodied* idea. From here, a conceptual exploration of functions and their possible properties begins. The effort of conceptual reorganization that Agustín undertakes is remarkable, and it prepares the ground for the in-depth research suggested by his work. With this arose a tangible need need to review the forms of existence of mathematical objects and what they meant. What was continuity in arithmetic reality? and How can we measure the size of a set? As mathematics pursued these questions, it became almost possible to abandon confidence in intuition completely, and this led to (among other consequences) a profound rupture: the dissociation of continuity from the differentiability of a function. All this awaits us below.

Keywords: Euclidean rigor, cognition, embodiment, completeness, continuity, cardinality

---

*lmorenoa@cinvestav.mx; corey.brady@smu.edu

The Mathematics Enthusiast, **ISSN 1551-3440, vol. 22, no. 3**, pp. 229-256.

Geometric intuition and the Pursuit of Rigor

Agustín proposed to abandon the method of *reasoning by analogy* (as practiced by Leonardo and Guillermo in passing from finite sums to infinite series); instead, he aimed to apply a standard of *Euclidean rigor*. In the opening of his text, *Cours d'Analyse*, he wrote:

> Regarding the methods, I have sought to give them all *the rigor that one requires in geometry*, so that one will never have to resort to reasons drawn from the generality of algebra (Cauchy, italics added).

Agustín thus sought to abandon the *unsafe* use of inductive methods in algebra, as Leonardo had employed them, and to rather *recover* Euclid's ideal that is crystallized in the *Elements of Geometry*. A key component of this effort is to re-construct analysis on the basis of a clear definition of limit. This definition of limit reads as follows:

> When the successive values attributed to a variable approach indefinitely a fixed value so as to end by differing from it by as little as one wishes, this last [fixed value] is called the limit of all the others. Thus, for example, an irrational number is the limit of diverse fractions which furnish more and more approximate values of it. (Edwards, 2012, p.310)

This is a dynamic, *embodied*, conception of limit. Just as Euclid was abstracting from objects and relations in the environment, Agustín was abstracting from motion and dynamic processes in the environment —this is what has been called *empirical abstraction*.

While this idea of a limit is grounded in the dynamism of motion or a process, the mathematical concept expands to describe situations that are not inherently dynamic processes. For example, *an irrational number can be defined as the limit of a sequence of rational numbers*. Underneath this statement (and making the definition valid) are camouflaged two central facts that will be identified later on, as crucial for the development, now arithmetically rigorous, of the analysis. These facts are (i) the *density of rational numbers* with respect the real numbers, and (ii) the *completeness property of the set of real numbers*. For Agustín, these properties are taken as given in advance. We feel it should be so in the teaching of Calculus, today.

Let us insist: Agustín's initial concern with the foundations of analysis was due to the somewhat uncertain state of analysis at the close of the 18th century. Functions, for example, were identified with analytic expressions. In fact, a function was understood to be given by a *single, closed* analytic expression. However, problems such as the study of a vibrating string were also key motivations and applications for analysis. The initial position of the string (at the moment of release) requires a description by more than one analytical expression.

Furthermore, it was understood that an analytical expression defined a *continuous* function. In fact, *uniqueness of the analytical expression defining a function was taken as synonymous with continuity*. But again, by stretching a string, a V-like shape was produced, which required *two* analytical expressions. Therefore, the initial state was not continuous, according to that previous notion since it was not described by a single analytical expression.

In addition to forcing the more general study of functions, modeling problems like the vibrating string raised the problem of analytical representation by means of trigonometric series of functions, and this entailed the *consideration of convergence*. That is, they again highlighted the need to *study the notion of limit, now at the level of series and sequences of functions*. Here it is necessary to distinguish *notion* from *concept*. Today, we would speak of the distinction between the *concept image* and the *concept definition* (Tall & Vinner, 1981).

This was the somewhat confused state Agustín found himself in during the first decades of his century. And in addition to the technical problems in analysis, a new social problem was looming before his eyes: commitment to education. All the work of his predecessors was guided by a deep intuition of mathematical correctness. However, their intuition was not, shall we say, *democratically distributed*. They shaped their findings for an audience of the few, for fellow initiates.

Therefore, Agustín made the decision to reorganize, on the basis of clear principles, the elements of analysis he inherited from his predecessors —Perhaps he was thinking of Leonardo…no one was like him. To put analysis in the hands of *ordinary people* (students, for instance), it was necessary to reconstruct the field..

Following his general description of the notion of limit, Agustín dealt with continuous functions. He abandoned the narrow conception of functions as defined by an analytical expression. And he defined the continuity of a function **f**, which is defined between two real numbers **a** and **b**, as follows:

> The function f(x) is continuous at x between the values a and b, if an infinitesimal increment of the variable x produces an infinitesimal increment of the function itself. In other words f(x+r)-f(x) is infinitesimal if r is infinitesimal.

It should be understood here that Agustín used the word *infinitesimal* to refer to a variable quantity whose limit is zero—that is, whose (positive) "numerical values…decrease indefinitely." Today, we would say that f is continuous at x if small changes in x produces small changes in the value of the function. This is a major departure from the previous definition of continuity in terms of the uniqueness of the analytical expression defining the function.

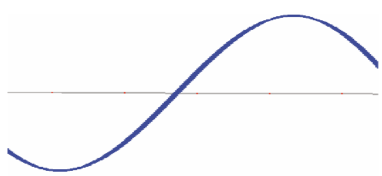
Next, motivated by the problem of guaranteeing the existence of the roots of a polynomial, Agustín tackled what would later be known as the *Intermediate Value Theorem*.

> If the function f(x) is continuous with respect to the variable x between A and B, and if we call C an intermediate value between f(A) and f(B), then the equation f(x) = C can be satisfied for at least one value of x between A and B.

Unlike his predecessors, Agustín was interested in properties common to families of functions (that is, functions that satisfy conditions—here, on values at endpoints of an interval) and not so much in particular functions. This theorem illustrates this fact. Illustrious predecessors such as Carlos Federico, when faced with the statement of this theorem, would have exclaimed: it is obvious, just look at the figure.

Figure 1

Looking for an Intermediate Value



The way to reason about this result is as follows: The curve having equation y = f(x) in the plane, and the line y = C must meet for some value of x between A and B. *This is evident if the hypotheses of the theorem are satisfied*, but it also highlights the way that "continuous" is being conceptualized. Agustín is using a *geometric* conception, and in this view, the power of the image is overwhelming. Perhaps this would have pleased Carlos Federico. We will say more about this theorem later on.

Thus, continuity, for Agustín, is the geometrical continuity provided by intuition, like the Euclidean conception of the line: a synthetic object. If that is what lives in the concept image of continuity, then one should not expect something radically different from what here is offered as a proof.

Carlos Federico similarly conceived of continuous functions as those whose graphs have properties associated with geometric continuity, in the sense that we can trace their graph *without lifting the pencil from the paper*. So it is an action, the movement of the hand, which corresponds to *continuity* of the function. But once the graph of the function is on the paper, our movement is frozen: The graph is a frozen object. This could have been the way to conceive of the line in Euclidean geometry. Consequently, the line became a *synthetic object.* If we conceive of the graph as what we see on the paper, then the Intermediate Value Theorem is obvious. If this were the case in Euclidean geometry, and we think it is, then it is not surprising that their conception of the basic objects was synthetic. Movement was absent in their mathematical epistemology—an epistemology of the frozen trace of movement.

Agustín's work can be described as a work of systematization based on the Euclidean continuum. It is not possible to attribute to him "deficiencies" such as those often attributed to Euclid when he is judged from today, that is, *from a conceptual framework that did not exist in his time*. We can close this section by reiterating that Agustín is to calculus what Euclid is to geometry. He objectified calculus as a field of mathematical exploration *per se*. That is, no longer to employ it purely as a tool for inquiry, but also to explore it *for its own sake* (which, for a time, coincided with the goals of striving to enable and improve the teaching of these ideas).

An (almost) rupture within mathematics

During the 19th century, beginning shortly after Agustín's work to revise analysis, mathematics went through a deep rupture, which unsettled the ways in which it had been conceived in the preceding centuries and shaped its subsequent development. One of these episodes was the creation of non-Euclidean geometries. Until then, mathematics had been based, almost exclusively, on the conviction that its results corresponded intimately with facts of the physical world. Therefore, what could be obtained from mathematics equated to truths about the natural world. Mathematicians affirmed that the natural world was not only a source of inspiration par excellence for mathematical creation but also legitimized their findings.

Non-Euclidean geometries shattered this conviction: the theorems of Euclidean geometry gradually lost their aura of absolute veracity and ended up merely constituting, like the theorems of other geometries, coherent possible *models* of space. Today it is difficult to imagine the disturbing effect that such a situation could have had on the mathematicians of the time.

In this climate, the casual attitude of 18th century mathematicians about convergence of infinite series raised red flags at the very heart of Calculus. And though the focus was the same, the style of the new critique was very different from Agustín's. After Euclidean Geometry had been unseated, neither Agustín's resort to Geometry as a standard nor his appeal to geometric intuitions as secure guides for thought were tenable.

The trigonometric series had already filled many researchers with perplexity, and they began to see a panorama of will-o-wisps, tempting mathematical travelers off the path of secure knowledge to their doom. One of them wrote that very few theorems of analysis had been proved *according to logical principles*. It is not necessary to add further testimonies to understand that in the very heart of mathematics the conviction—that mathematics was as consistent as nature itself—had been shattered and another view was beginning to take shape: the *validity* of results can be guaranteed by logical rigor, but not the *veracity*. This is inevitable when mathematical structures are seen only as a plausible model: a model is a *map*, not the modeled *territory* itself.

That is the way things were when Ricardo began teaching analysis. In his own presentation of the field, at a certain point he had to use the fact that *an increasing and bounded sequence is convergent*. He wrote that in teaching this result, *he was forced to resort to geometric evidence*. And he added that resorting to geometric intuition in a first course is very useful from the didactic point of view. However, he wrote that this intuitive presentation of the aforementioned result cannot be said to be *scientific.* Here was a different perspective from Agustín's, where the geometric and the rigorous were allied. For Ricardo, geometric reasoning was a scaffold for learners toward a more "scientific" endpoint. In his book *Essays on the Theory of Numbers,* Ricardo wrote that he thought continuity based in geometric intuition*, was not providing the rigorous approach needed* for developing a scientific foundation for differential calculus. In contrast, his work established that **completeness**, that is, continuity from an *arithmetic* point of view, was the DNA of the theory of analysis.

There are several equivalent versions of completeness. It is important to remember that when Agustín introduced his definition of the limit of a sequence, he added that, for example, π can be obtained as the limit of a sequence of approximations by *rational* numbers. That every irrational could be obtained as the limit of a sequence of rational approximations, was something that Agustín took for granted. We recall here the sentence: *we do not know what we see*—which Agustín had before his eyes—*rather, we see what we know*... and Agustín knew it. He knew that he could obtain an irrational as the limit of a succession of rational numbers, following the path of decimal representations: 3, 3.1, 3.14, 3.1415, 3.14159, ...

His assertion thus depended on the (guaranteed) existence of the decimal representations of the real numbers. That is, although we may not know how to generate the decimal expansion to an arbitrary degree of precision, we know that this expansion exists. *Here we are venturing—without going any further now—into a major theme of the epistemology of mathematics: the forms of existence of its objects.*

It has been explained that the *episteme* of a generation, of its culture, is constituted by ideas that penetrate so deeply in the collective thought, that they define to a great extent both *how* one thinks and *about what* one thinks in those historical moments. Agustín thought of the continuity of the line in synthetic terms and that object, the line, was *colonized* by the numbers that were represented (principally) by the positional decimal system. Between two numbers, let us say, between π and 3.2, there was room for an infinity of rational numbers. For example, 3.15, 3.19, and so on. Agustín thus had the mechanism to use the completeness of the arithmetized line, but he was not fully aware of what it meant. He did not use the words "completeness" nor "dense." But the inquisitive mind of Ricardo understood that in Agustín's assertion was hidden the secret of continuity *from a numerical/arithmetic point of view*. The idea of decimal representation could be translated as *approaching with destination.* This is what we are trying to say with the expression: *an increasing and bounded sequence is convergent*. Of course, it too embodies a dynamic way of thinking.

Let us insist: the most common statement of the completeness of the real numbers is that *the set of all real numbers is exactly the set of all decimal representations*. Agustín took this assertion as valid, and yet, as we have already mentioned, this was not an explicit decision on his part.

Ricardo was aware of this fact, we believe, because the intellectual atmosphere of his time has changed. Now, mathematics itself was already an object of study in itself.

Ricardo knew that if every real number has a decimal expansion, finite or infinite, then he could prove that a bounded increasing sequence is convergent. But his inquisitive mind led him to ask *why* one has to accept that any number has a decimal representation. He did not want to take it as something given beforehand; rather, he set out to do something almost heroic: to *construct* the irrational numbers from the rational ones. We will not discuss here the fine details of his original construction. Instead, just the idea which was its starting point.

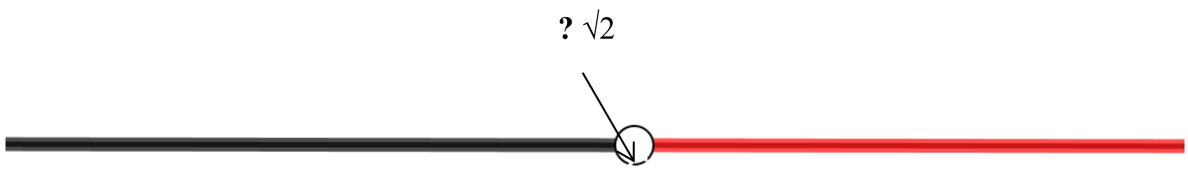
Here we return to an idea already discussed, but with an additional twist. Let's imagine the following: Ricardo knew that there are irrational numbers such as √2. Consider the set:

A={1, 1.4, 1.41, 1.4142…}

As we move forward in this sequence, we approach √2 with improved accuracy. If there were only rational numbers on the line, then this sequence would lead to a "hole" on the line instead of the number √2. This number would be missing! Therefore, arithmetic continuity (completeness) is equivalent to the *absence of holes* in the line when we *colonize* it with all the real numbers.

Figure 2

Is There a Number There?



First Agustín, with his eagerness to systematize the concept of limit, and then Ricardo, recognizing the need to make explicit the transition from geometric continuity to arithmetic continuity, opened the way to a new mathematical age. In a sense, Ricardo´s work relied on Agustín´s. What would have happened if they had met? Imagine the dialogue between them!

As we have seen above in his comment about the pedagogical utility of geometric intuition, Ricardo makes explicit in a very few lines that mathematics and mathematics education--though they have different needs, can find a conciliatory ground. It would be desirable if today there were the same degree of sensitivity to the problems of teaching and learning *about* the problems of mathematics. The didactics of mathematics is a discipline that does not have to subordinate its lines of research to the lines of mathematics itself. However, we hasten to say, we do not ignore the intimate link between both fields of inquiry. Human cognition and the epistemology of mathematics are permanent participants in these educational reflections.

We should try to communicate to our students the permanent attitude of developing deep thinking about the basic, elementary ideas of mathematics. These ideas are elementary in the sense that they are the building blocks out of which a plastic structure is assembled. As made clear by

René Thom in 1972, in the teaching of mathematics it is the *development of meaning* that should be privileged, rather than rigor. (Here, he means *artificial rigor*.)

Felix, a younger thinker, was very interested in the work of both Agustín and Ricardo. A permanent concern of his was to find a pedagogical bridge between elementary mathematics as taught in schools and university mathematics. He was aware that there was a *discontinuity* between these two educational levels. He wrote and lectured extensively on this issue. He was immersed in a cultural atmosphere of high mathematical density, and he deployed what today is called the *pedagogical content knowledge* of the mathematics teacher. Felix recalls that Carlos Federico, although he belonged to a generation that had gradually introduced a more critical spirit into mathematical work, used his intuition of space as the basis for his demonstrations. For Felix, there was a permanent tension between the new arithmetical rigor and intuitive thinking. His views are based on his epistemic concerns and his permanent involvement in the education of teachers, above all. Let us quote him fully:

> I do not grant that the arithmetized science is the essence of mathematics; and my remarks have therefore the two-fold character of positive approbation, and negative disapproval. For since I consider that the essential point is not the mere putting of the argument into the arithmetical form, but the more rigid logic obtained by means of this form, it seems to me desirable—and this is the positive side of my thesis—to subject the remaining divisions of mathematics to a fresh investigation based on the arithmetical foundation of analysis. On the other hand, I have to point out most emphatically—and this is the negative part of my task—that it is not possible to treat mathematics exhaustively by the method of logical deduction alone, but that, even at the present time, *intuition has its special province*.

The movement, towards a version of mathematics with *higher levels of mathematical rigor*, was imminent. Mathematics was entering a new age. In addition to the inductive method, the deductive method claimed its place. But Felix recognized that *neither would yield in favor of the other*. A tension that had not been so intense since the times of Greek Euclidean geometry and more recently, since the advent of non-Euclidean geometries, was becoming tangible. There was not one voice in mathematics, but two.

A new perspective and some proofs

Accepting that every real number has a decimal representation, we will offer a *possible* proof in Ricardo's hands of the proposition:

An increasing and bounded sequence is convergent.

Let us imagine that Ricardo took the existence of the decimal representation as a given and then explored its consequences. He would have proceeded as follows:

Let (x_n) be an increasing and bounded sequence. Let us suppose, without losing generality of the argument, that the elements of the sequence are positive. Let us write these elements:

$x_1=N_1.\, a_{11}a_{12}a_{13}\ldots$
$x_2=N_2.a_{21}a_{22}a_{23}\ldots$
…..
$x_n=N_n.a_{n1}a_{n2}a_{n3}a_{n4}\ldots$. and so on.

The numbers N_1, N_2, N_3, … are the integer parts of each element of the original sequence (x_n). As the sequence is increasing and bounded, the integer parts have values from a finite set of integer values (if B is an integer upper bound, this set is a subset of the values $\{N_1, N_1+1, \ldots, B\}$). The sequence of the N_k thus has a largest value. That is, from some element of the sequence (xn) on, the sequence of integer parts becomes constant. Let M be the largest (and constant) value of the N_k.

Now let us consider the sequence formed by the first decimal places of each x_n: $(a_{11}, a_{21}, \ldots, a_{n1}, \ldots)$. As the original sequence is increasing, then this last sequence has to be increasing from some a_{rk} on (after the element for which the integer parts become constant). Moreover, *this* sequence also has to be bounded, and as it is a sequence of (one-digit) natural numbers, it must have a maximum value M_1.

Along the same line of argument, let M_2 denote the maximum value reached by the sequence of the second decimal places $(a_{12}, a_{22}, a_{32},\ldots)$. Now it is clear how to continue:

The limit of the sequence is denoted $M.M_1M_2M_3\ldots$and the proof is complete. This is the *magic* of decimal representation!

Ricardo knew that completeness was *the fundamental proposition needed for making calculus/analysis scientific*. Now he could continue his teaching, knowing that the intuitive hypothesis he used to explain convergence, had a *scientific* foundation.

Let us go a step further: Staying at the intuitive level, he could have used this example to explore the possible pathway that *could lead* to the proof: It was known that *e=2.718281828459045…*

Let us write the first terms of the sequence converging to e: $1+\sum_{k=1}^{\infty}\frac{1}{k!}$

1) 2.000…(corresponding to n=1)
2) 2.5000…
3) 2.6666…
4) 2.708333…
5) 2.716666…
6) 2.718055…
7) 2.71853…
8) 2.718277…

In this partial list, one can detect from which term onwards, the sequence of the first digits reaches its maximum value and remains constant afterwards—similarly with the subsequent digits. For instance, the digit in the first decimal place (7) reaches its maximum (and constant) value when n=4. The second digit reaches its maximum when n=5. The reader can continue!

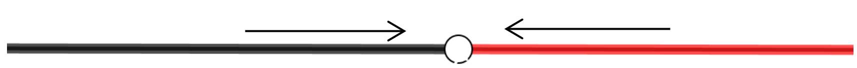
Jorge, a good friend of Ricardo's, was very interested in these mathematical subjects. He observed that if you consider a sequence of *nested* intervals:

$$[a_1, b_1] \supseteq [a_2, b_2] \supseteq [a_3, b_3] \supseteq \ldots \supseteq [a_n, b_n] \supseteq \ldots$$

…such that the length $b_n - a_n \to 0$, then there is exactly **one** point (number) **x** common to all the members of the nested intervals. This follows from the fact that the sequence $a_1, a_2, a_3, \ldots, a_n, \ldots$ is increasing and bounded, and so convergent. Let A be the limit of this sequence. On the other hand, the sequence $b_1, b_2, b_3, \ldots, b_n, \ldots$ is decreasing and bounded, so it, too, is convergent. Let B be the limit of this sequence. By hypothesis $b_n - a_n \to 0$, and therefore A=B=**x.**

Figure 3

Capturing the Number



We can imagine the sequence $a_1, a_2, a_3, \ldots, a_n, \ldots$ is "walking" from the left (black) side of the line towards **x,** and the sequence $b_1, b_2, b_3, \ldots, b_n, \ldots$, from the right towards **x** . This is an *embodied way of understanding* the two convergence processes.

Here is an example to shed more clarity on this proposition: Consider the number π = 3.1415…. This decimal expansion also gives us a condensed way to construct a nested sequence of intervals collapsing to π. For instance:

$[3, 4] \supseteq [3.1, 3.2] \supseteq [3.14, 3.15] \supseteq [3.141, 3.142] \supseteq [3.1415, 3.1416] \supseteq \ldots$

The number π is the unique element in the intersection of this sequence of nested intervals.

Probably you have already realized the profound relations between the proposition on the convergence of a (decreasing or) increasing and bounded sequence and the principle of nested intervals. It is not by chance: they are equivalent. That is, if you accept one of them, you can obtain the other as a consequence.

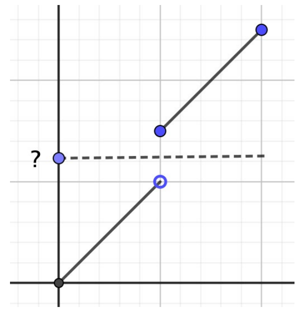
In the story we are telling you about the arithmetization of Calculus, we take as a starting point that every real number has a decimal representation. This seems the simplest way to establish the completeness property, either in the form of increasing (decreasing) bounded sequences, or by means of nested intervals. What we have told is intended to give a bit more of clarity to a fact that, even for Agustín, was so intuitive: *the number line has no holes*! That is the idea we try to capture the significance of the fact that the number line, the real number line, is *complete*.

At the beginning of this story, we mentioned that the *Intermediate Value Theorem* was one of the basic results about continuous functions. We promised that we would come back to it. Well, *now we are ready*. The subject of *continuous functions* is delicate but fundamental for the study of arithmetically rigorous versions of Calculus. Informally, we say that a function is continuous at a point b with f(b)=B, if f takes on values close to B, for all values in its domain close to b. That is, informally, if a is close to b, then f(a) is close to f(b).

For a long time, Agustín´s colleagues thought that continuity was equivalent to the condition that the function took on all intermediate values between two values, say a and b. This *is* equivalent to continuity whenever the continuous function is increasing or decreasing on [a,b], because omitting a value between f(a) and f(b) would mean the function has “jumped.” Graphically, this would look like:

Figure 4

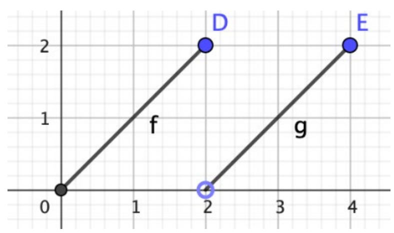
Monotonicity and Intermediate Values



Now, let's see how a function that is not strictly increasing or decreasing can take all values between two given values, without being continuous:

Figure 5

The Intermediate Value Property Does Not Imply Continuity



The domain of this function is the interval [0, 4]; f(0)=0 and f(4)=2; and the function takes all values between 0 and 2. However, the function is *discontinuous* at x=2. The reader can think of other examples. One might ask: How is it possible that these Agustín and his colleagues did not think of an example as simple as the one we have just presented? The reason is that in the background of their work lurked the conception that a function *should have only one analytical representation* and not two like the example we have just presented. That example, for those thinkers, represented two functions, not one. Though Agustín was working to extend the calculus theorems of his contemporaries (and himself, of course) to more general functions, the conceptual transition was not complete.

The Intermediate Value Theorem expresses that continuity guarantees the "intermediate value" property, while the above shows that the converse is not true.

Without further ado, let us consider how to prove the Intermediate Value Theorem. The following statement is equivalent to the Theorem, substituting convenient values:

> Let f be a continuous function defined on the interval [0, 1]. If f(0) and f(1) have opposite signs, one positive, the other negative, then there must exist c between 0 and 1 such that f(c) = 0.

Proof. We can assume that $f(0)<0$ and $f(1)>0$. We are aiming to locate c such that $f(c)=0$. We bisect the interval [0, 1] to obtain [0, 1/2] and [1/2, 1]. If $f(1/2)=0$ we are done. Otherwise, we know that $f(1/2)<0$ or $f(1/2)>0$. Suppose that $f(1/2)>0$. Then we hunt for c "to the left" of ½. We bisect the interval [0, 1/2] and consider the value f(1/4). If it is 0, we are done. If $f(1/4)\neq0$ we observe the sign of f(1/4). If it is positive, we again "look to the left" and select the interval [0, 1/4]; otherwise, f(1/4) is negative, and we select [1/4, 1/2]. We continue this process, generating a nested sequence of intervals $[a_i, b_i]$ for which $f(a_i) < 0$ and $f(b_i) > 0$.

Again, our process is as follows: at each step, we bisect the interval of focus, starting with [0, 1], and we thus generate a nested sequence of intervals whose lengths tend to zero. Therefore, they determine a single point c between 0 and 1. Now, we argue, f(c) must equal 0.

What would happen if not? Suppose $f(c)=C>0$. Since the function f is continuous, then for domain values close to c f takes on values close to C. So, since C is strictly positive there must be a small open interval (m, n) around c where the function remains positive. But the lengths of the intervals of the sequence of nested intervals above tended to zero, so after some point, they must all have been contained in the interval (m, n). This is a contradiction, because we constructed the intervals so that the function took on different signs at the endpoints, and yet we have asserted that it is positive throughout the interval (m, n), which contains them.

The same contradiction would arise if $f(c)<0$. Thus $f(c)=0$, ending the proof. **NOTE.** We can easily transform any closed interval [a, b] into [0, 1]. That is why this statement is equivalent to the full Intermediate Value Theorem.

We can see through these results of Ricardo and Jorge, which were in turn based on the pioneering work of Agustín, how the discipline was transforming its intuitive arguments into arguments with increasing arithmetic weight.

In the teaching from that point on, this trend translated into an eagerness to introduce (prematurely, in our opinion) an arithmetized version of the ideas we have discussed in this text. Without affirming that the organization of the knowledge presented here should follow a chronological order according to its historical development (which we will not do), it is convenient to return to Felix's suggestion that *intuition has its special province*.

How intuitive-informal or how arithmetical-rigorous should we make the presentation of a subject of study? It is a very delicate didactical question. Perhaps nobody has the answer. We can only aspire to find partial and contextual responses. One thing is clear however: we need both. When we lean too much to the side of rigor, it is good for us to consider the reflection of a disciple of Felix, Jaime, who wrote:

> The analysis of to-day is indeed a transparent science. Built up on the simple notion of number, its truths are the most solidly established in the whole range of human knowledge. It is, however, not to be overlooked that the price paid for this clearness is appalling, it is total separation from the world of our senses.

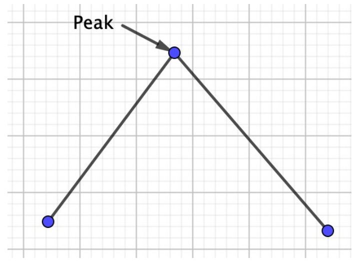
Lest we consider the arithmetical approach to functions to yield only cold rigor or dry results, our friend Felix has asked us to discuss a few other rich and generative mathematical ideas that follow from this way of thinking. Let us not forget the importance of Tedoro and his profound impact on these ideas. Being, as he is, a master in the art of mathematics and possessing profound pedagogical ideas, we have found it convenient to follow his suggestions. Our purpose in this essay is not to develop a complete history of the transition from calculus to analysis. We propose something more modest: to leave along the way some guidelines for those who venture later to explore a territory that is not yet known to them.

Following Felix's suggestion, then, let us introduce his friend Teodoro. In his time, he was well known for his profound mathematical and pedagogical work. In his younger years he had been a primary school teacher, and this strengthened his teaching inclination. He introduced a genuinely arithmetical rigor into mathematical analysis by formulating fundamental concepts with a precision hitherto unknown.

Teodoro was particularly pleased to have shown that continuous functions did not necessarily possess derivatives. In fact, he constructed an example of a continuous function which at no point in its domain possessed a derivative. This overturned a long-held conviction about continuous functions, namely that in the *worst case*, they would fail to have a derivative at *isolated* points. For example, this might happen wherever the graph of the function had a “peak”.

Figure 6

Loss of Differentiability



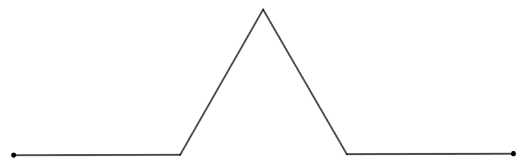
Yet Teodoro constructed a function that was continuous and yet failed to have a derivative at *any* point. The consequences of this discovery were profound. Just as the advent of hyperbolic geometry overturned the belief in Euclidean geometry as an *exact model* of physical space, mathematical “monsters” (cf. Lakatos, 1976) such as continuous functions without derivatives called into question mathematical intuition as the sole compass for the discovery and justification of both past and new results.

However, that did not dethrone intuition completely. Felix, who felt the impact of the new ways of advancing mathematical work, repeatedly insisted that *intuition had its province*. It seems to us that he was thinking above all about ways of teaching, about how school curricula were going to be transformed and how this might affect the relationship between school and university. The tension between intuitive approaches on the one hand and rigorous-arithmetical approaches on the other hand seemed to be unavoidable. *We hold the conviction that this tension is inherent to our cognitive nature*. We will return to this subject later.

When Teodoro's example of the continuous function without derivative became known, many scholars expressed their dissatisfaction with what they considered to have gone too far in the direction of arithmetic rigor. They saw that example as palpable evidence of that excess. One of them, whom we will call Koch, proposed a construction, which he called geometric, of a continuous curve without tangent lines. This construction was inspired by Teodoro's example but aimed to make the phenomenon more accessible. First, let us describe Koch's construction. It starts with a segment, say, of length 1. Then the central third is extracted and replaced by two segments that would form an equilateral triangle with the extracted segment. It thus increases the length of the total path by a ratio of 4:3. The figure below is created:

Figure 7

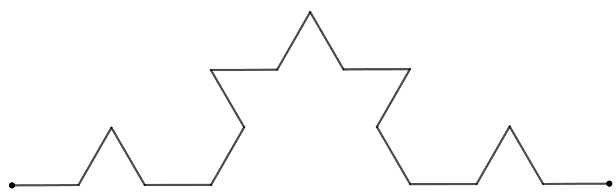
Spreading Non-Differentiability



Next, the procedure from the previous step can be repeated on each of the four segments of the figure. We obtain the following figure:

Figure 8

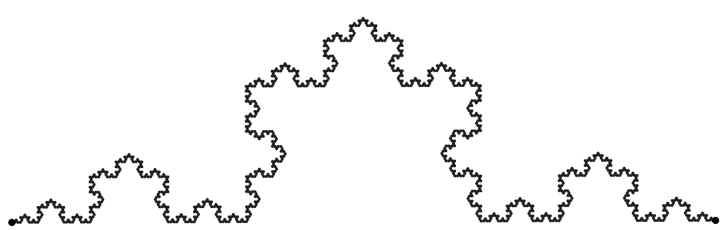
Even More Non-Differentiability



Now we know how to continue: on each of the segments we repeat the initial construction; after *five steps* we obtain the following result:

Figure 9

Non-Differentiability Everywhere



The reader may make an effort to imagine the result, if we continue this process *indefinitely*. Let's anticipate the answer: we will have a continuous "curve" that has no tangent at any point. But...wait a minute! There is a lurking paradox. On one hand, we know that it does not have a tangent at a point because it has a peak there. But, the very idea of a "peak" includes the idea of two sides, two very small, straight sides. On each of those straight sides, there *is* a tangent that coincides with the respective side. Which is it—peaks everywhere or sides everywhere? We have seen students raising and struggling with this paradox, as they attempt to understand the curve.

Before leaving this topic, let's consider a similarly accessible construction by a Japanese colleague named Takagi that actually yields a *function*. Tagaki actually discovered an infinite *family* of functions, but we'll focus on two of them.

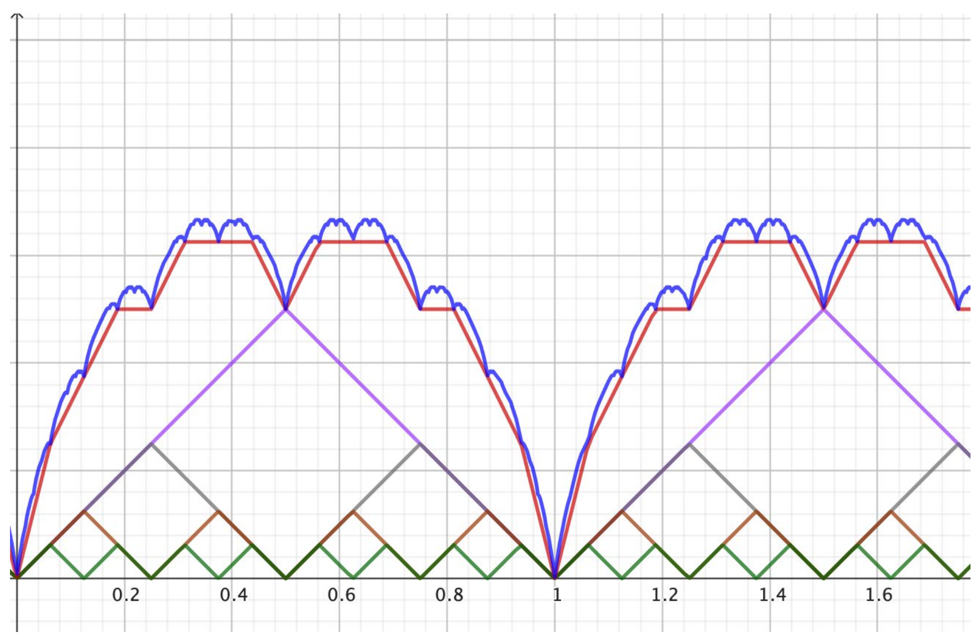
Both of Takagi's functions that we will consider follow the same construction approach: they are built by adding together "sawtooth" functions. Like the Koch curve, each of these functions has peaks and sides. We will build them so that when they have more peaks, each peak is less tall. And our function will be the sum of an infinite series of these sawtooth functions.

Here is the first of Takagi's functions, which British mathematicians called the "blancmange," after a desert pudding with a similar shape. In building this function, at each step we will add a sawtooth that has double the 'frequency' of the prior one, and half the 'amplitude' of that function.

The first figure below shows the first four sawtooth functions, along with their sum (in red), that is, the fourth partial blancmange sum. We also have shown (in blue) the *ninth* partial blancmange sum, which helps to show the function that is emerging.

Figure 10

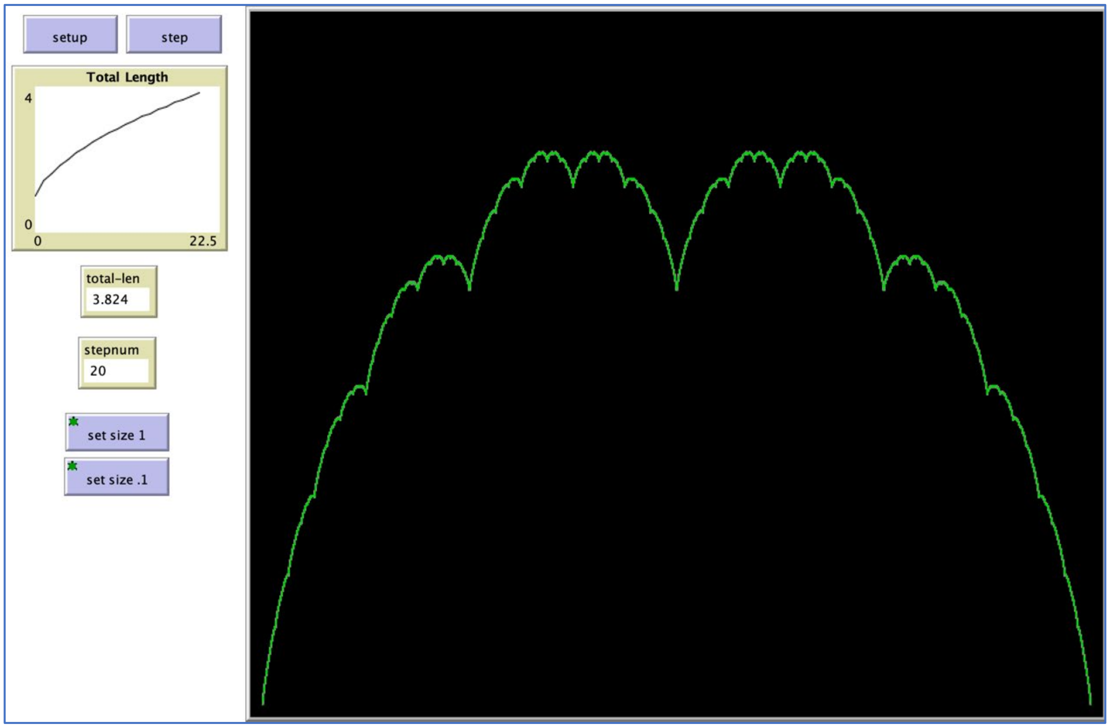
Fourth Iteration of the Blancmange Function, the 'Sawtooth' Functions that Compose It; and the Ninth Iteration for Reference.



Next, here is a representation of the 20th partial blancmange sum function:

Figure 11

Twentieth Iteration of the Blancmange Function.



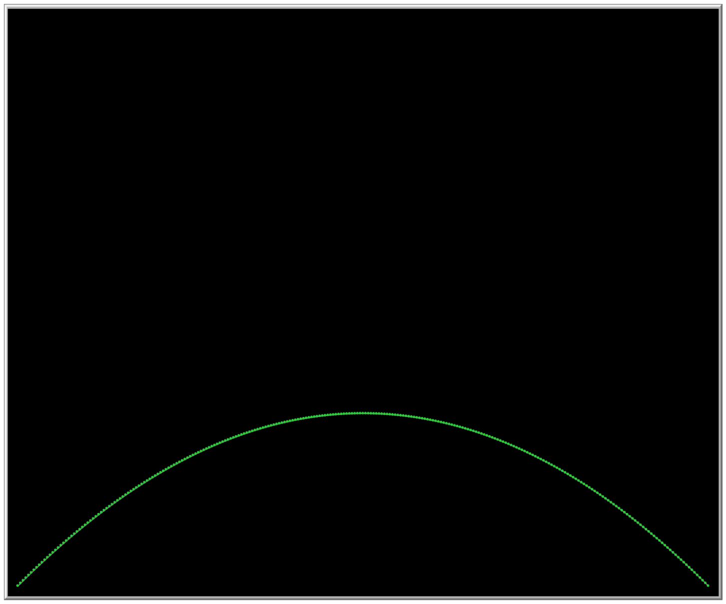
The blancmange function, created by the infinite sum of the sawtooth functions, has the property that it is *nowhere differentiable*. This connects it to the Koch curve, which fails to have a tangent line at every point.

Because the sawtooth functions are very simple in nature, we can start to reason about points on the Blancmange function. Looking at the partial-sum functions, we notice jagged points that are local minimum points. These appear at zeroes of the sawtooth functions. Because of how the sawtooths work, when an x-value is a zero of one sawtooth, it is a zero of *all of the subsequent ones*! So, these valley-points are actually points on the final blancmange function—for these x-values, the y-value will *never* grow any higher. However, for x-values *nearby* these valley-points, subsequent sawtooths *will* add to their y-value, which will make the valleys more and more steep. A peak of *any* kind is a point of non-differentiability, but these peaks are in a sense *infinitely* steep.

So, this blancmange function really *is* a monster! But it may become even more shocking when we consider its "sibling." For this one, we will use the exact same construction procedure, but at each step, we will multiply the amplitude of the next sawtooth we add by ¼ rather than ½. As if by magic, the monstrosity of the blancmange becomes a parabola:

Figure 12

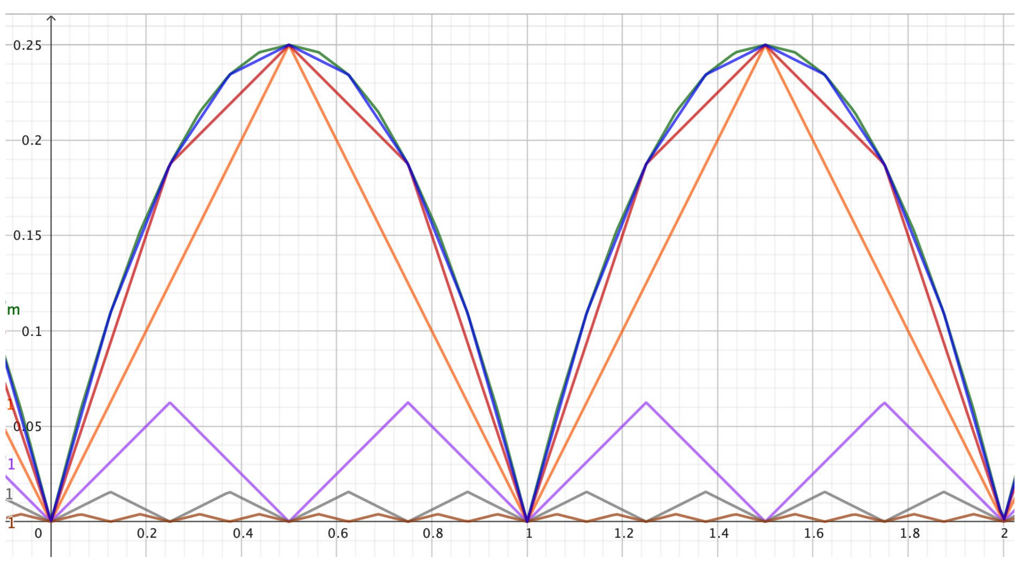
With One small Change, the Blancmange Procedure Generates a Parabola



With a 4:1 zoom, here are the first four component functions and partial sums:

Figure 13

Sawtooth Functions that Sum to Produce the Parabola.



Sereno in fact knew of this construction. The picture helps a modern audience to draw connections between his work with circles and with parabolas (Recall, Sereno filled the circle by adding isosceles triangles to the sides of inscribed polygons. This construction is not so different!) And now, in the context of Takagi's work, it becomes clear that the parabola is also very close to the blancmange: monsters are lurking close by the sunny meadow where we have built our mathematics.

As in the case of irrational numbers, the Koch curve, and the blancmange function, the "last iteration of the process" is an ideal object. It is an abstraction. The figures we can draw are steps of an iterative process, but they show the result of a finite number of steps, in an *infinite* process. The Koch curve is *the limit* of this process. It is an ideal object. We mention the irrational numbers because they offer a similar situation. If we think of the decimal representation of an irrational...we cannot write it completely! In fact, it has an infinity of digits that in a certain sense *emerge randomly*. While we may be unable write *all* the digits of some rational numbers (for example, 2.345345345...), apart from its integer part 2, the next three digits 345 are repeated indefinitely. So, we know everything about that number. The amount of information that it contains is finite.

In contrast, for an irrational (e.g., π), because there is no explicit rule to produce successive digits, they appear randomly, and the quantity of information it contains is infinite. In fact, a number such as π is the *embodiment of randomness*.

Mathematics is confronted, every moment, with the tension between the *concrete* and the *abstract*. Koch´s and Teodoro´s curves are palpable examples of this dialectic between the concrete and the abstract, as also are the irrational numbers. These reflections inevitably refer us back to the didactics of mathematics.

Following René Thom, a central concern of learning consists in coming to understand how mathematical objects exist and how their meanings can be developed. A mathematical object conceals in itself a multiplicity of meanings. In short, it is polysemic. How then to bring a student's embodied experience closer to the more rigorous formulations—particularly when these formulations seem to hide or dislodge embodied reasoning? We could say without exaggerating that the didactics of mathematics finds in this question a primary motivation for its development. Didactics has been sensible to the obstructions to learning that students repeatedly encounter when dealing with formalized versions of mathematical concepts and propositions that ignore the complex processes involved in *developing them*. We cannot ignore the fact that what is *logically* simple (or what can be formulated in logically simple terms) does not necessarily coincide with what is *cognitively* simple. The logical structure of a theorem, for example, lives within its cognitive structure but does not exhaust it. As we write these lines, we are reminded of what Felix said at the time: *intuition has its province*. It is consistent with this premise, that the structuring of knowledge gives form and new mathematical substance to ideas that are born in embodied interactions and experiences. Pointing out with an arrow ($\rightarrow$) the process of convergence of a sequence awakens in the student a sense of a dynamic process which then, as the idea enters the formalized space of arithmetic, passes to a static medium. This kind of evocation of embodied ideas, followed by their transfer to a formalized medium requires sustained attention, without which the rupture between the process of learning and the object of knowledge becomes more profound.

Until this point in the development of mathematics, infinity appeared only as *potential infinity*. That is, the term *infinity* was applied to describe a process that went on and on indefinitely, without ending.

It has been said that a distinctive feature of humans is that we can predict the outcome of a process when it is still unfinished. This ability has allowed us, especially in mathematics, to *transform a process into an object*. The expression 3.14159...indicates the first stages of an infinite approximation process which we identify with an ideal object—the number π. In general, when we talk about an irrational number, we conceive of it as a number that has an “infinity of decimal digits”...that is to say it is "something"—always incomplete but, in our mind, a well-defined object. Similarly, when we speak of a circle as the limit of a succession of inscribed polygons whose number of sides increases and increases “to infinity,” which is where we “find” the circle. Clearly, the idea of limit as the reification or “object-ification” of a dynamic process is key.

The development of mathematics reveals a need to distinguish between this—what is called *potential* infinity, a process that continues indefinitely—and *actual* infinity, which refers to a totality. For instance, consider the *set* of natural numbers. If we say 3/7 is not a natural number, in the end we are saying that the collection of natural numbers has an identity of its own. This set is infinite – it is not *becoming* infinite; rather, it has an infinite *cardinality*, or “count.”

There are some interesting passages in Galileo's book *Discourses and Mathematical Demonstrations Relating to Two New Sciences* where actual infinity appears as an undeveloped concept. The book is written as a dialogue among three friends. A passage where they are considering the natural numbers is worth quoting in detail:

> **Salviati**: If I should ask further how many squares [i.e. square numbers, 1, 4, 9, 16, …] there are one might reply truly that there are as many as the corresponding number of roots, since every square has its own root and every root its own square, while no square has more than one root and no root more than one square.

And he continues:

> **Salviati**: But if I inquire how many roots there are, it cannot be denied that there are as many as there are numbers because every number is a root of some square…
> Yet at the outset we said there are many more numbers than squares, since the larger portion of them are not squares. Not only so, but the proportionate number of squares diminishes as we pass to larger numbers.
>
> **Sagredo**. What then must one conclude under these circumstances?
>
> **Salviati**. So far as I see we can only infer that the totality of all numbers is infinite, that the number of squares is infinite…neither is the number of squares less than the totality of all numbers, nor the latter greater than the former; and finally, the attributes "equal," "greater," and "less," are not applicable to infinity.

We have selected these brief passages because they clearly show that Galileo was *close* to affirming that natural numbers and their squares have the same "size," but that he turned away from this assertion, with the statement that infinite “sets” cannot be compared in this way. We will return to these lines a little further on.

Jorge began his research on mathematical infinity motivated by the structure of sets of real numbers. For instance, he studied functions that had different behaviors on the part of their domain formed by rational numbers and on other parts, formed by irrationals. Thus, these functions did not have a uniform behavior. Over time, he experimented with different types of subset in the domain of his functions, trying to determine characteristics of a *minimum subset* that would cause the same behavior to apply over the entire domain of the function.

Searching for an answer to this question, Jorge developed a kind of *surgery* on the subsets of the real numbers and from there, gradually, emerged his study of *the size of abstract sets*. The first problem he faced was to define what should be understood by *number of elements of a set.* This led to the study of mathematical infinity. Then, under Ricardo´s suggestion, Jorge used the idea of *one-to-one correspondence between two sets*. For instance, imagine you enter a theater, and you notice that no chairs are empty, and no one is standing. Your conclusion is: there are as many chairs as there are people—in other words, the same *number* of chairs and people. Importantly, you could do this *without having to count* either chairs or people!

Galileo came close to understanding the problem in these terms. It seems to us that by initially considering infinite sets, the old idea that *the whole is greater than its parts* appeared before him as an epistemological obstacle —something was missing in the conceptual structure of the mathematical culture in which he lived. He was aware that he could *count* the squares and that meant the squares constituted an infinite set *potentially*. But he could not indicate the *number of elements* with any precise language, as in the case of finite sets. This became an obstruction and remained so for centuries. Until Jorge's work, only infinity, as a process, that is *potential* infinity, had a place in mathematics. Here is where the crucial idea appears: to count the number of elements of an infinite set, the appropriate instrument is the 1-1 function. The first step in the task of assessing the size of an infinite set is to compare with the set of natural numbers. Given two sets, A and B, if you can define a 1-1 function from A to B, then we say that A has the same number of elements as a subset of B (the image of A). If the function is also surjective, we will say that A and B have an equal number of elements, or that they have the same *cardinality*.

Definitions are crucial in mathematics. Once you have defined the cardinality of a set, you have to subordinate your thinking to the definition. That does not mean abandoning intuition, but it does mean the rules of the game have changed. Definitions introduce precision into your thinking, and they also guide inquiry, for good or ill.

With the *cardinality* of a set defined, Jorge began exploring the size of sets. For instance, natural numbers and even numbers. The bijection $n \leftrightarrow 2n$ proves that the even numbers have the same cardinality as the natural numbers (as a whole set!). That is, according to the definition of number of elements, there are as many naturals as even numbers. The set of odd natural numbers has the same cardinality as the set as the set of even numbers.

Another category of surprise appears when we compare the naturals with the *rational* numbers. Our previous experience with this set is that it is *dense* in the set of real numbers: There is a rational number between any two other numbers on the real line. Intuitively, the conjecture must be that the set of rational number is *larger* than the set of natural numbers. But let us remember that we must follow to the letter the definition by means of 1-1 functions. If we can prove, contrary to the intuitions we bring to the problem, that the set of rational numbers has the same cardinal number as the set of natural numbers, then, we must accept both infinite sets have the same size. Here we are witnessing a tension between our intuitions and our *formal reasoning,* which is being guided, in this case, by the definition of cardinality.

Below, we give an idea of how to prove that rational numbers have the same cardinality as natural numbers. We use the word *countable* to indicate that a set has the same cardinality as the set of natural numbers or is finite. The even numbers, the odds, and the prime numbers are all examples of countable sets. Now: How can we prove that the set of rational numbers is countable?

Consider the following lists: A_1, A_2, A_3, A_4, …, A_k, with k a natural number:

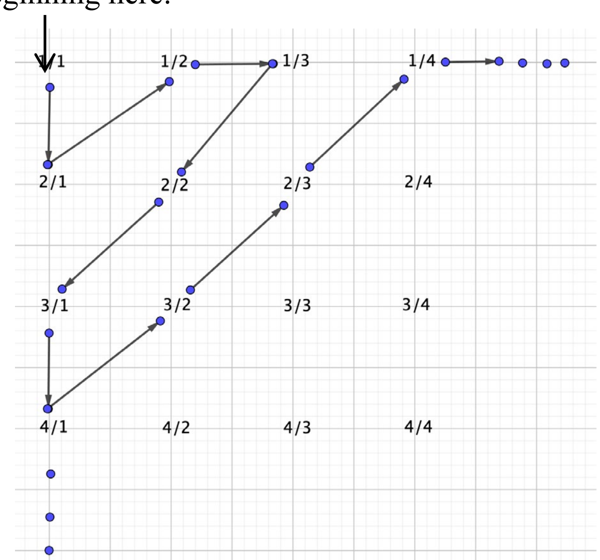
A_1 = {1/1 2/1, 3/1, 4/1, … }
A_2 = {1/2, 2/2, 3/2, 4/2, … }
A_3 = {1/3, 2/3, 3/3, 4/3, … }
A_4 = {1/4, 2/4, 3/4, 4/4, … }.

We can produce a single list by following the arrows as indicated below:

Figure 14

Listing the Rationals

Beginning here:



The new total list is: {1/1, 2/1, 1/2, 1/3, 2/2, 3/1, 4/1, 3/2, 2/3, 1/4, …}. By identifying a way of iterating through the collection of all A_k elements, we have created a function mapping the natural numbers to this collection. That is, the collection is countable. It is true that we have only generated the list of the positive rational numbers, but the reader can prove that the union of the positive and negative rational numbers is countable (the approach is the same for the positive and negative whole numbers).

Proving that the rational numbers are countable came as a surprise, but this surprise was minor compared to the next result: **the real numbers are NOT countable.** This was a colossal theorem. From here, Jorge began a mathematical journey totally unknown to the generations that preceded him.

To say that the reals are not countable means that no matter how large a *listing* of real numbers is, this listing will always leave out some real number.

Let us prove what seems to be an even more aggressive statement—namely that the subset of real numbers in the interval [0,1] is not countable. We will suppose they *are* countable and produce a contradiction. If they are countable, then we can write a list such as this:

$x_1 = 0.a_{11}a_{12}a_{13}a_{14}\ldots$
$x_2 = 0.a_{21}a_{22}a_{23}a_{24}\ldots$
$x_3 = 0.a_{31}a_{32}a_{33}a_{34}\ldots,$

…and so forth, which enumerates all real numbers in the interval.

Now, to produce a contradiction, let's construct a number B between 0 and 1 that is *not* in this list. We will call it $0.b_1b_2b_3b_4$.... and we will define it like this: $b_n = 1$ if $a_{nn} \neq 1$, and $b_n = 2$ if $a_{nn}=1$. So, our number B differs from the first real in the list in the tenths place; it differs from the second in the list in the hundredths place; and so forth.

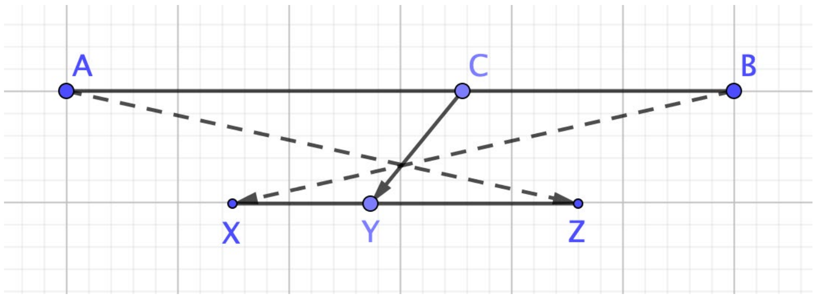
This means that B is not in this list because for every element of the list, B differs from that number in some decimal place: $b_1 \neq a_{11}$, $b_2 \neq a_{22}$, and so forth. Because of the way the number B has been generated, this method of proof is called the *diagonal* method.

The importance of this result is immense. Imagine all infinite sets had the same cardinality. Then, we would essentially have one example. With this result, we know we have at least two different infinite sets, let us say N, the natural numbers (model for all countable sets) and the real numbers, our **first** model for uncountable sets. (We say *first* because from here, Jorge was able to go on to discover an *infinite number of infinities, all different from each other.*)

What could happen if instead of the interval between 0 and 1 we had chosen another interval? The answer is: nothing essentially different would have happened. Let us illustrate this fact:

Figure 15

Measuring Non-Enumerability



Point A corresponds to Z and point B corresponds to X. Through the intersection point of the segments AZ and BX, we can make C correspond to Y; C is any point in segment AB. This correspondence defines a 1-1 and surjective function from the segment AB onto segment XZ. This means that these segments have the same cardinality. Amazing! These are the first and basic steps of the theory of mathematical infinity.

As we said before, Jorge had to face strong criticisms of his work. However, some brilliant mathematicians came to his rescue. One of them, a very special one, David, wrote: *No one will expel us from the paradise that Jorge has created for us*.

CODA

Currently, mathematics shows an even more axiomatic and formal face. It would seem that it has distanced itself from the world of applications and that it is only interested in its internal developments. No doubt, one can get this impression if one looks at this huge discipline without studying the processes of mathematical learning and discovery: what do these continuous functions without derivatives, or the so-called fractals, have to do with human experience or the natural world? In the recent past these objects were considered “monsters,” or pathological constructions, but today it has become clear that far from being perverse curiosities, these strange objects have served to model levels of material (and eventually social) reality that had not been possible with the mathematical tools of the past. Even the most abstruse objects have been finding applications outside the field of mathematics itself. This demonstrates that even the most sustained efforts to detach ourselves from the external world by elaborating advanced theories have not been successful; and the reason, it seems to us, is anchored to a fundamental fact: our hybrid cognitive nature, which is articulated around an interplay between intuitive thinking and analytical thinking–

–in other words, around our holistic and symbolic cognitive capacities which are coextensive with our environment.

We cannot detach ourselves from that environment: this is a profoundly evolutionary fact. A brilliant scientist, Freeman Dyson, wrote:

> The mathematicians who created the monsters regarded them as important in showing that the world of pure mathematics contains a richness of possibilities going far beyond the simple structures that they saw in nature. Twentieth century mathematics flowered in the belief that it had transcended completely the limitations imposed by its natural origins.
>
> Now…we see that nature has played a joke on the mathematicians. The 19th-century mathematicians may have been lacking in imagination, but nature was not. The same pathological structures that the mathematicians invented to break loose from 19th-century naturalism turn out to be inherent in familiar objects all around us in nature (Dyson, 1978).

References

Cauchy, A. L. (2009) *Cauchy's Cours d'Analyse: An Annotated Translation.* (R. E. Bradley & C. E. Sandifer, Trans.). New York, NY: Springer. (Original work published 1821).

Bottazzini, U. (1986). *The higher calculus: A history of real and complex analysis from Euler to Weierstrass*. New York: Springer.

Dauben, J.W. (1990). *Georg Cantor: His Mathematics and Philosophy of Infinite*. Princeton: Princeton University Press.

Davis, P., & Hersh, R. (1986). *The Mathematical Experience*. Boston: Birkhäuser.

Dyson, F. (1978). Characterizing Irregularity. [Review of the book Fractals—Form, Chance, and Dimension, by B. B. Mandelbrot]. *Science, 200*(4342), 677-678.

Edwards, C. H. (2012). T*he historical development of the calculus*. New York: Springer Science & Business Media.

Lakatos, I. (1976). *Proofs and refutations.* Cambridge, UK: Cambridge University Press.

Tall, D., & Vinner, S. (1981). Concept image and concept definition in mathematics with particular reference to limits and continuity. *Educational Studies in Mathematics*, *12*(2), 151-169.