

Inner simplicity vs. outer simplicity

Étienne Ghys

Editors' note: The text is an edited transcript of the author's conference talk.

For me, mathematics is just about understanding. And understanding is a personal and private feeling. However, to appreciate and express this feeling, you need to communicate with others—you need to use language. So there are necessarily two aspects in mathematics: one is very personal, emotional and internal, and the other is more public and external. Today I want to express this very naïve idea for mathematicians that we should distinguish between two kinds of simplicities. Something could be very simple for me, in my mind, and in my way of knowing mathematics, and yet be very difficult to articulate or write down in a mathematical paper. And conversely, something can be very easy to write down or say in just one sentence of English or French or whatever, and nevertheless be all but completely inaccessible to my mind. This basic distinction is something which I believe to be classical but, nevertheless, we mathematicians conflate the two. We keep forgetting that writing mathematics is not the same as understanding mathematics.

Let me begin with a memory that I have from when I was a student a long time ago. I was reading a book by a very famous French mathematician, Jean-Pierre Serre. Here is the cover of the book (Figure 1). For many years I was convinced that the title of the book was a joke. How else, I wondered, can these algebras be complex and simple at the same time? For mathematicians, of course, the words ‘complex’ and ‘semisimple’ have totally different meanings than their everyday ones. ‘Complex’ means complex number and ‘semisimple’ means a sum of simple objects. So, for many, many years, I was convinced that this was a joke. Recently, actually one year ago, I had the opportunity to speak with Jean-Pierre Serre, this very, very famous mathematician, who is now 85 years old. I dared ask him the question: “is this a joke?” With sincere curiosity, he replied, “What? Why is it funny?” He never noticed the apparent contradiction. It was not a joke

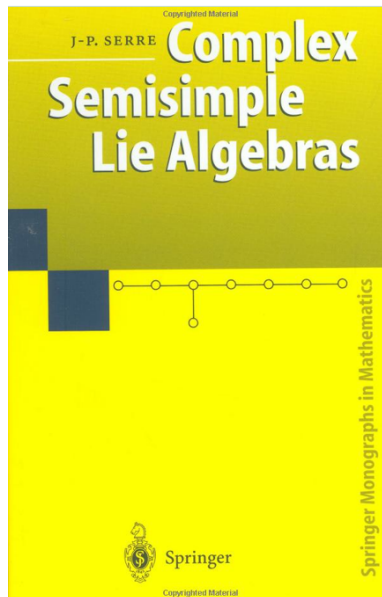


Figure 1: “Why is it funny?”

to him. Mathematicians use words as words, and they don’t want to use the words with their meaning.

There is a famous quote attributed to David Hilbert that says you can replace all the words in mathematics arbitrarily. Instead of ‘line,’ you could say ‘chair,’ and instead of ‘point,’ you could say ‘bottle,”and then you could say that “in between two bottles, there is one chair,” and the mathematics would be unchanged. This is the point of view of Hilbert, which is not at all my point of view.

So this is the first aspect, that there is in mathematics an external simplicity which is conveyed by the language, and this language is somewhat artificial—it is made out of words which are not fully subject to meaning. Oversimplifying the picture, one could distinguish these two aspects by saying that on the logic side there is Hilbert writing words without looking for meanings for these words, while Poincaré is on the intuition side (Figure 2).

Notice here that, and this is my favorite part, the latter image is from a chocolate bar wrapper. Poincaré was so famous they would use his photograph on chocolates. (Do you know one mathematician today whose picture could sell chocolate?) Hilbert was basically focused on transmitting mathematics, and Poincaré was focused on understanding mathematics. This is

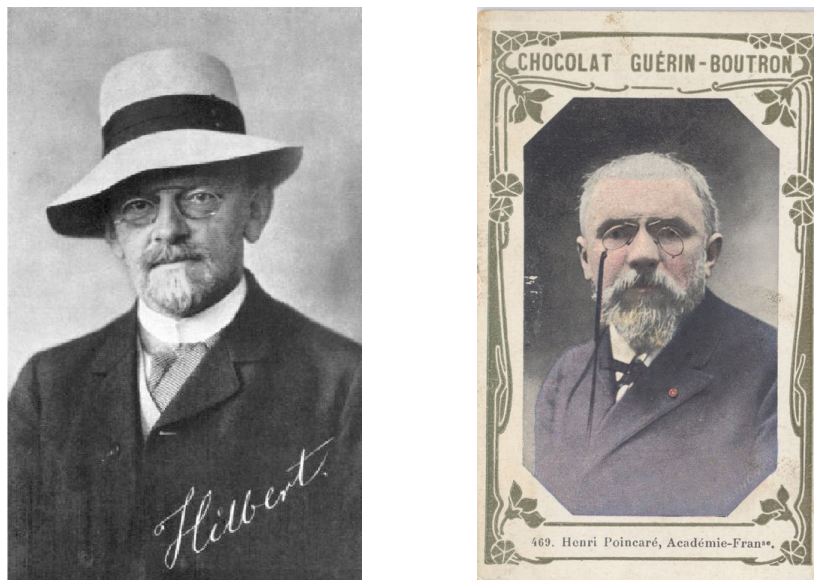


Figure 2: Logic vs. Intuition

one way that I want to distinguish between inner and outer simplicity.

Before we start, since I am the first speaker, I thought it could be a good idea to open the dictionary at the words ‘simplicity’ and ‘complexity’.

simplicity (n.) late 14c., from Old French *simplicity* (French *simplicité*), from Latin *simplicitatem* (nominative *simplicitas*) “state of being simple.” from *simplex* (genitive *simplices*) “simple”.

simplex (adj.) “characterized by a single part,” 1590s, from Latin *simplex* “single, simple” from PIE root **sem-* “one, together” (cf. Latin *semper* “always,” literally “once for all;” Sanskrit *sam* “together;” see same) + **plac-* “-fold.” The noun is attested from 1892.

complex (adj.) 1650s, “composed of parts,” from French *complexe* “complicated, complex, intricate” (17c.), from Latin *complexus* “surrounding, encompassing,” past participle of *complecti* “to encircle, embrace,” in transferred use, “to hold fast, master, comprehend,” from *com-* “with” (see *com-*) + *plectere* “to weave, braid, twine, entwine,” from PIE **plek-to-*, from root **plek-* “to plait” (see *ply*). The meaning “not easily analyzed” is first recorded 1715. *Complex sentence* is attested from 1881.

This is perhaps obvious, especially to such a scholarly and learned audience

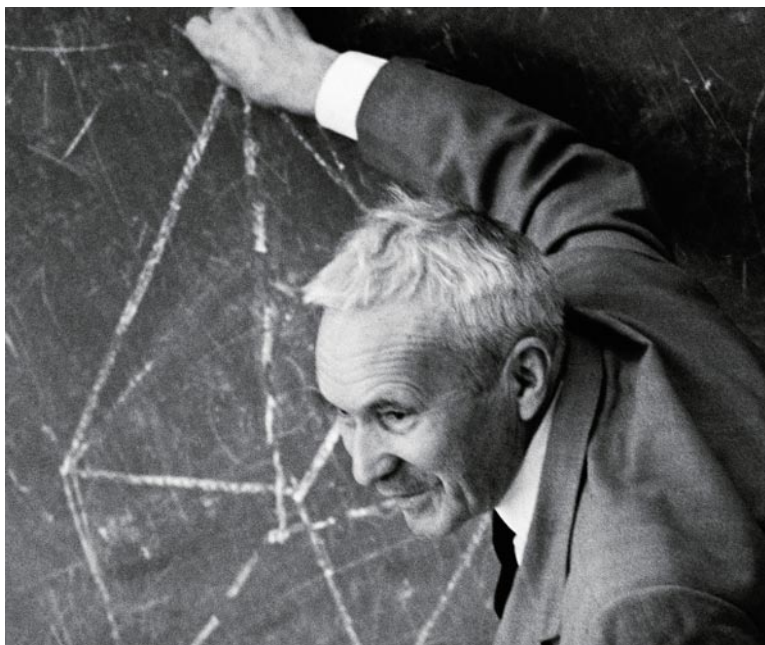


Figure 3: Andrey Kolmogorov

as I have here today, but I would add that it may not be as obvious for you as it is for French speaking people. The word ‘simple’ comes from the French word *plier*, “to fold.” Something simple is folded only once, and it’s complex when it has many folds. (The closest cognates in English might be the verbs “ply” and “plait.”) To explain something is to ‘unfold it.’ Complexity and simplicity are related to folding in all directions, and this is something we will keep in mind.

Let’s begin with outer simplicity. Given its reliance on words, there is an obvious measure of complexity here: the so-called Kolmogorov complexity. In the 1960s, Andrey Kolmogorov (Figure 3) had the idea of defining complexity of something to be the length of the shortest explanation of that something. By merely asking how many words are needed to describe something, you get a measure of the complexity of this object.

Complexity = Length of the shortest description

Here is an example, a simple example. Take this painting by Kazimir Malevich from 1915 (Figure 4). I can describe it to you in, let’s say, five

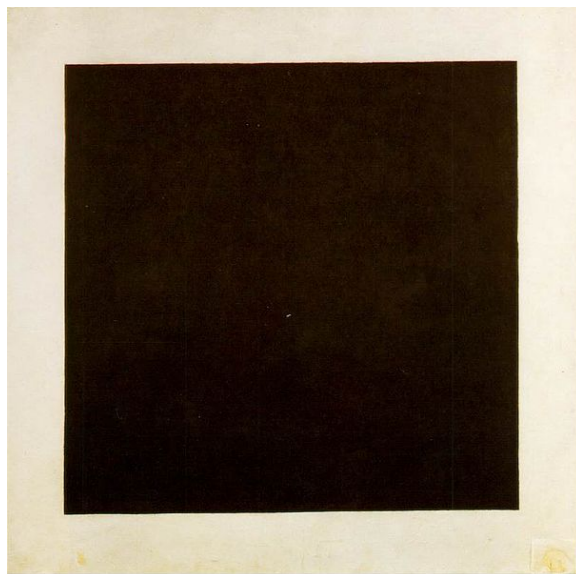


Figure 4: *Black Square*, Kazimir Malevich (1915)

or six sentences: it's a square with such size, and it's white, and inside it there is a smaller square which is black. I could give the precise blackness and whiteness of the two squares. So this is a very simple object. That was Malevich, let me show you my own art object (Figure 5). This is a totally random object. It's a square, and in the square there are many dots. I asked my computer to put yellow or orange dots here, but it's totally random. If you ask me to describe it to you in detail, the only way that I can do it is to describe it *dot by dot*. I will need a very, very long sentence that might begin "the first point is yellow, the second point is red," etc. It will be a very long description. So, in Kolmogorov's terminology, this is a complex object, and Malevich's is a simple object.

Here is the third object—one that is very famous, at least in the mathematical realm—the Mandelbrot set (Figure 6). It looks complicated and, mathematically, it is complicated. But for Kolmogorov it's a very simple object. In order to produce this picture, it may take a computer a long time, days, or weeks, or more, but the computer program that describes the Mandelbrot set is two lines long. So, from the Kolmogorov's point of view, this object is very simple. This is the first concept of simplicity, outside simplicity, the length of what you need to describe it. Clearly, it is not

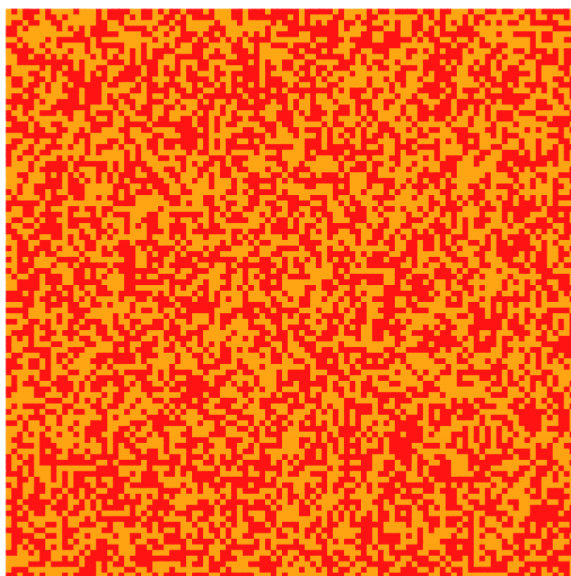


Figure 5: High Kolmogorov complexity

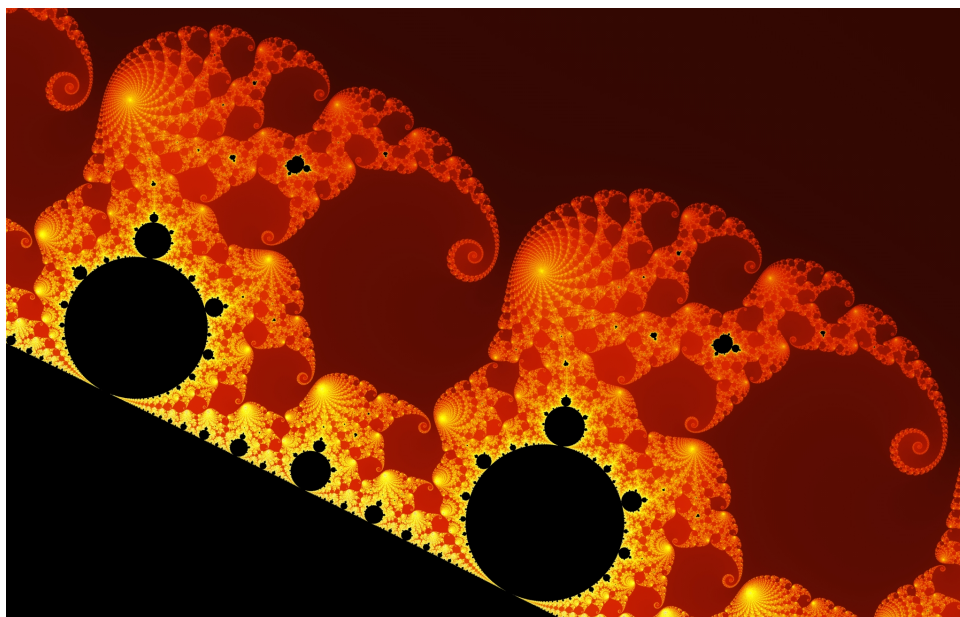


Figure 6: Low Kolmogorov complexity

CHAPITRE I

CORPS FINIS

Tous les corps considérés ci-dessous sont supposés commutatifs.

§ 1. Généralités

1.1. Corps premiers et corps finis.

L'intersection des sous-corps d'un corps K en est le plus petit sous-corps; il contient l'image canonique de \mathbf{Z} , isomorphe en tant qu'anneau intègre à \mathbf{Z} ou à $\mathbf{Z}/p\mathbf{Z}$ avec p premier; il est donc isomorphe, soit à \mathbf{Q} , soit au corps $\mathbf{Z}/p\mathbf{Z}$.

Figure 7: *Cours d'arithmétique* by Jean-Pierre Serre, page 1

satisfactory. I mean, for me, I don't want to consider the Mandelbrot set as being something simple. This object is complicated for me. It is made out of many folds.

Let me give you another example, a personal example, of a simple linguistic thing that is complicated. Or, at least, it was complicated to me when I was a student. Again, I will take the example from Jean-Pierre Serre. Serre wrote a wonderful book for students on number theory called *Cours d'arithmétique*. I opened it when I was, I think, 19. Here is what I found on the first page (Figure 7). The first sentence of the book begins, "L'intersection...". (I'll explain in a moment why I am showing this in French.) I can tell you that I spent two days on this one sentence. It's only one sentence, but looking back at this sentence, I see now that it is just perfect. There is nothing to change in it; every single word, even the smallest, is important in its own way. I wanted to show you the English translation, but the English translation is so bad compared to the French of Jean-pierre Serre. Serre's language is so efficient, so elegant, so simple. It is so simple that I don't understand it. Even the smallest words, like "d'un corps K en", this two-letter word "en" is fundamental. Everything, every single word is fundamental. Yet, from the Kolmogorov point of view, this is very simple. But as a student I knew almost nothing about "anneaux

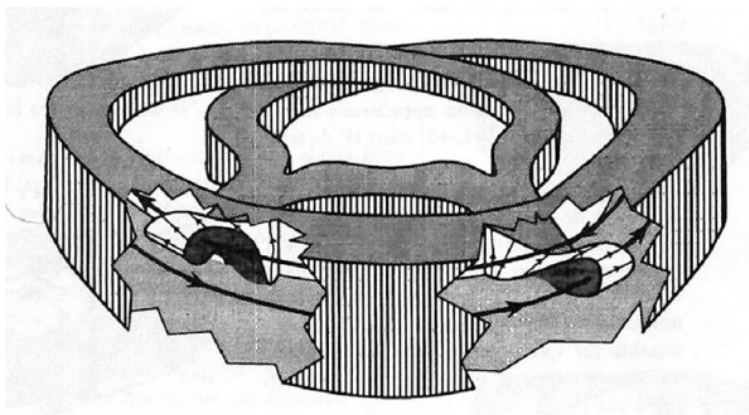


Figure 8: An illustration

intègres” and all these other things. It looked so complicated. Finally, at the end of the second day, all of a sudden, I grasped it and I was so happy that I could understand it. From Kolmogorov’s point of view, it’s simple, and yet for me—and, I imagine many students—it’s not simple.

Let me give you another example from Jean-Pierre Serre. I should mention that Serre is perhaps the most famous French mathematician. We mathematicians from France, we consider him to be some kind of (semi) God. He writes exquisitely. Most of my students, when they are writing their PhD theses, or whenever they write badly, which is usually the case, I say to them, “go to the library, open any book of Jean-Pierre Serre, and try to copy!” In terms of elegance and economy, there is nothing better. Back to the example I wanted to mention. A long time ago, maybe fifteen years ago, I was giving a talk in the Bourbaki seminar. I was describing a construction in dynamical systems due to Krystyna Kuperberg of a very fascinating counterexample to an old conjecture of Herbert Seifert (the construction of a vector field on the 3-sphere with no periodic orbits). This is a wonderful, simple idea, really wonderful. For my talk, I prepared pictures, and here is one of the pictures that I showed (Figure 8). It’s not important to my point that you understand what this object is. In my talk I explained the construction saying, you know, “you do this, and this...” [*gesturing towards the picture with both hands*]. After the talk, well, I thought it was successful, people were happy. Then Jean-Pierre Serre came up to me and said, “That was interesting what you said. I have a question.” And he asked, “Would you consider this to be a theorem?” In other words, he was questioning

whether the fact that I was using pictures, and not words, didn't disqualify me from transmitting mathematics. My feeling, and this feeling is shared by others who you will see in a second, is that pictures and, more than pictures, even movies, should be incorporated into the world of mathematics as genuine tools of proof. Not just for fun, but for veracity, and for presenting mathematics.

So let me explain something to show that I'm far from being the only one to think this way. We'll discuss Hilbert's twenty-fourth problem in this meeting, but today I want to discuss the zeroth Hilbert problem. When Hilbert gave his famous lecture in Paris on problems for the future of mathematics, his paper contained twenty-three problems. These were preceded by a general introduction on what makes a good problem, what is interesting, where should we go, etc. There is something in this introduction that I want to show you because I believe that, to this day, it presents a fundamental question for mathematics. The point is that we should incorporate pictures as genuine tools for understanding and transmitting mathematics. So, here's an extract from Hilbert's introduction of what I call his zeroth problem:

"To new concepts correspond, necessarily, new signs. These we choose in such a way that they remind us of the phenomena which were the occasion for the formation of the new concepts. So the geometrical figures are signs or mnemonic symbols of space intuition and are used as such by all mathematicians. Who does not always use along with with the double inequality $a > b > c$ the picture of three points following one another on a straight line as the geometrical picture of the idea of 'between'? Who does not make use of drawings of segments and rectangle enclosed in one another, when it is required to prove with perfect rigor a difficult theorem on the continuity of functions or the existence of points of condensation? Who could dispense with the figure of the triangle, the circle with its center, or with the cross of three perpendicular axes? Or who would give up the representation of the vector field, or the picture of a family of curves or surfaces with its envelope which plays so important a part in differential geometry, in the theory of differential equations, in the foundation of the calculus of variations and in other purely mathematical sciences?

"The arithmetical symbols are written diagrams and the geometrical figures are graphic formulas; and no mathematician could spare these graphic formulas, any more than in calculation the insertion and removal of parentheses or the use of other analytic signs.

"The use of geometrical signs as a means of strict proof presupposes the exact knowledge and complete mastery of the axioms which un-

derlie those figures; and in order that these geometrical figures may be incorporated in the general treasure of mathematical signs, there is necessary a rigorous axiomatic investigation of their conceptual content. Just as in adding two numbers, one must place the digits under each other in the right order, so that only the rules of calculation, i.e., the axioms of arithmetic, determine the correct use of the digits, so the use of geometrical signs is determined by the axioms of geometrical concepts and their combinations.”

So Hilbert is asking for a language of pictures, for ways of presenting mathematics simply while not restricted to the use of letters and languages. Recently I had a discussion with some choreographers, and they face a similar problem. They are looking for a notation for dance. How would you denote choreography? They have several ways of doing it, for example, one is called Benesh Movement Notation, but there are many other possibilities. And they have exactly the same problem: why should we restrict ourselves to a linear, totally ordered language in order to describe mathematics, since we are not linearly ordered in our mind? Or at least I am not.

“The agreement between geometrical and arithmetical thought is shown also in that we do not habitually follow the chain of reasoning back to the axioms in arithmetical, any more than in geometrical discussions. On the contrary we apply, especially in first attacking a problem, a rapid, unconscious, not absolutely sure combination, trusting to a certain arithmetical feeling for the behavior of the arithmetical symbols, which we could dispense with as little in arithmetic as with the geometrical imagination in geometry. As an example of an arithmetical theory operating rigorously with geometrical ideas and signs, I may mention Minkowski’s work, *Die Geometrie der Zahlen*.”

This is not the usual way that we think of Hilbert; here he is praying for a better use of pictures. Now let us get back to Poincaré. Poincaré was not at all motivated by words or language, which are on the outside of mathematics. He was motivated by the inside of mathematics, by intuition. He warns that we should not compare mathematics with the game of chess. Anyone can easily learn the rules of the game, you can check if a game is fulfilling the rules, but it is clear that you are not a mathematician if you only know the rules of the game. You need to have some global understanding of the subject, and from that point of view logic is totally useless. Here’s what he writes:

“If you are present at a game of chess, it will not suffice, for the understanding of the game, to know the rules for moving the pieces. That will only enable you to recognize that each move has been made conformably to these rules, and this knowledge will truly have very little value. Yet this is what the reader of a book on mathematics would do if he were a logician only. To understand the game is wholly another matter; it is to know why the player moves this piece rather than that other which he could have moved without breaking the rules of the game. It is to perceive the inward reason which makes of this series of successive moves a sort of organized whole. This faculty is still more necessary for the player himself, that is, for the inventor.”

This reminds me of something. When you use your smart phone to look up direction with Google Maps, it’s amazing how quickly it finds the best path from A to B. Basically, this is what we are trying to do in mathematics. We want to go somewhere, and we are looking for the best path. I don’t know if any of you have looked at the algorithm that Google Maps uses. It’s called the A* algorithm. It’s a very, very clever way of finding your way in an unknown country, and I strongly suggest that you take a look at this remarkable algorithm. Maybe it could be used, by analogy, to understand better how mathematicians work, how sometimes we try to move forward by first moving backward so as to change course.

What I want to say is that at present the connection between mathematics on the outside and the inside is not good. We should improve it. We should write mathematics in a different way. Hilbert is suggesting that we should use pictures, I would even add that we should use movies.

Let me give you an example. There is this theorem of Stephen Smale that implies that, in a particular sense, it is possible to turn the 2-sphere inside out (we can evert it in 3-space). This is not an easy theorem to prove, the proof is formal and difficult to understand. However, about 20 years ago, Silvio Levy, Delle Maxwell and Tamara Munzner created a movie on this result called *Outside In*, which is based on ideas of Bill Thurston. This 22-minute long movie uses extraordinary computer graphics to show you how the eversion works. Of course, it still would not qualify as a proof, but it comes very close to a proof. And if we follow the advice of Hilbert, we ought to devise the rules with which to transform such a film into a genuine proof. In the future, maybe tomorrow, or in ten years, one should be able to publish proofs using movies, as soon as they are certified by some certification that we do not as yet know how to do.

Another thing that I consider important is that we should think of the

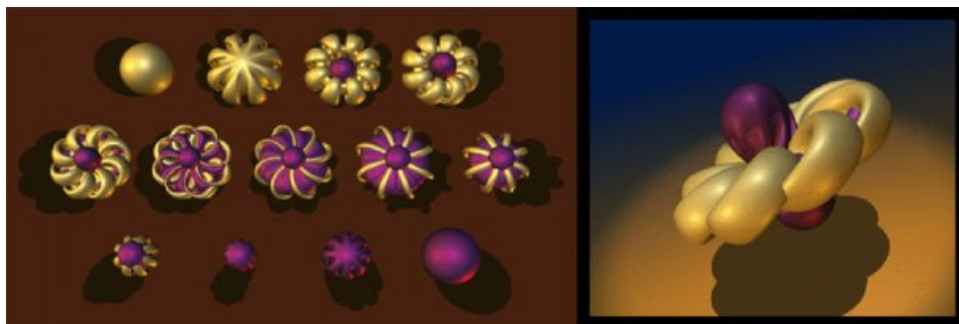


Figure 9: Outside in

way we write mathematics. For many years we have written papers from A to Z, and it is well-known that no mathematician would open the paper or book and start by the beginning and go to the end. We go forward, we jump, we come back, we go to some other place. So we should be able to write mathematics (and not only mathematics, actually) in a non-linear way. Today's technologies—computers, e-books, the Internet—make it possible to do. So why don't we do that? I think it's time to create papers that are not just standard papers going from A to Z.

An exciting possibility, at least for me, in this direction will begin next week, when I will meet with a group of eighteen mathematicians who plan to write a new book on algebraic topology for graduate students. We want to do it this way. We don't want to write a book with pages. We want to write a book which is completely electronic, in which you can travel in a way which is adapted to you as a reader. Of course, this requires some planning before we start. But we feel that we have to try to adapt the outer ways of describing mathematics to the inner ways of our readers.

Here is a very different idea about mathematical writing. This is this crazy idea that comes from Paul Erdős that somewhere in heaven there is THE BOOK, and in it are some jewels, some wonderful proofs, and we should work toward these beautiful proofs, simple proofs, elegant proofs. Martin Aigner and Günter Ziegler's book *Proofs from THE BOOK* (Figure 10) is supposed to contain a few hundred of those jewels. I'm not really convinced. I don't know how many of you read this book, but some of the theorems inside this book are really wonderful. But I tried myself to read it, and I can guarantee that in most proofs, not all but most, they are just wonderful. And then you close the book, and let's see, now it's one year later, I have forgotten them.

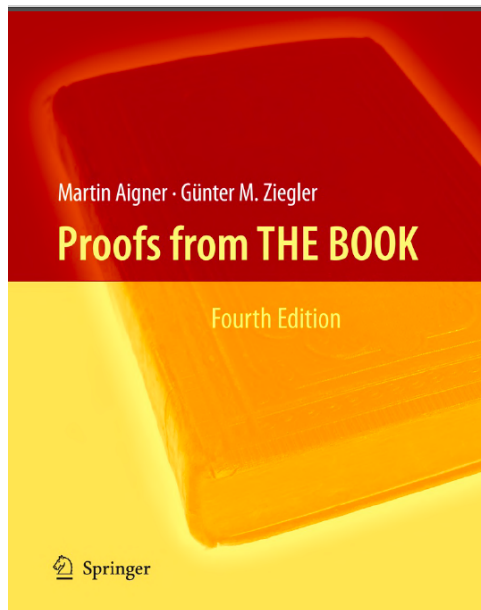


Figure 10: *Proofs from THE BOOK* by Aigner and Ziegler

This is a bad sign. I mean, when I understand something, by definition, I don't forget it. The concept of beauty here, in my opinion, cannot be the correct one. Simplicity is not the correct one.

I would like to finish by explaining some mathematics. I'm not sure if it's true in this country as well, but when I was a student, I was told that you should never, never, never give a talk, without stating a theorem. So I decided that maybe I could spend the last minutes of this talk mentioning a theorem that I understand and that I think I will never forget. I will not give you the proof, but I will explain this theorem because I believe it reveals something about the way the brain understands mathematics. I'm not a neurobiologist, I know even less about psychology, but I think it's something fundamental. It's a simple, fundamental idea, and I want to share it with you.

What I'd like to try is to discuss how we can understand large networks. To begin, here is a picture of a network of neurons (Figure 11). You have hundreds of thousands of neurons, and they are connected in a way that you don't really understand. And you want to describe this structure. Do you know how many neurons I have in my brain? [*Audience member: "A hundred*

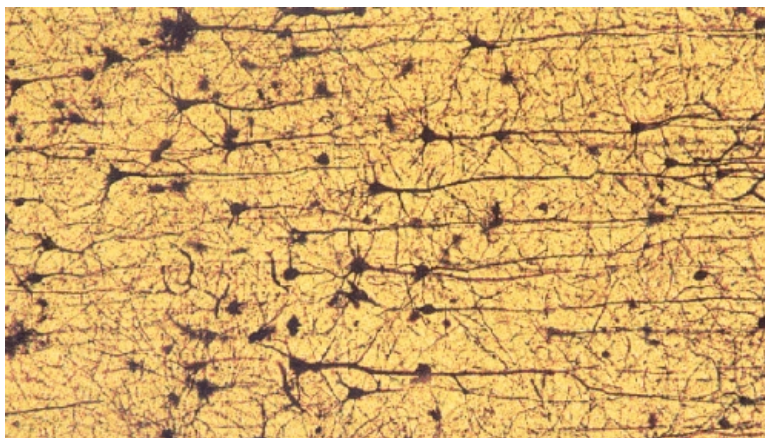


Figure 11: Large networks

billion.”] I think I have only ten billion in my brain. Okay I think it’s ten billion. For comparison, here’s a picture, albeit somewhat vague, of the Internet (Figure 12). How many HTML pages are there in the world? Ten billion. So the number of pages in the Internet is approximately the same as the number of neurons in your brain. The difference is in the connectivity of these two networks. A typical webpage is connected to about twenty other webpages. But a typical neuron is connected to ten thousand neurons, so things are much more connected in my brain. Another main difference is that communication is much faster on the Internet than in my brain. This is because Internet connections use electricity or light, while communications in my brain use biological or chemical reactions, which are much slower. So my brain is slower but better connected, and the Internet is less connected but faster.

How can we understand these two huge structures? This is part of the motivation for the theorem I want to mention. It’s a theorem of Endre Szemerédi (Figure 13) called the Szemerédi regularity theorem. As you will see, it’s a very general theorem that is true for all networks. It conveys the idea that all networks, no matter how big they are, can be understood in finite terms, so to speak. Let me explain. Here’s a network (Figure 14). A network is just a bunch of points, which could be whatever you want, and some of them are connected by links, or edges, which you draw between the points as in the picture. Of course, this picture is reasonable because the number of dots used is small. It’s totally impossible for me to draw a picture

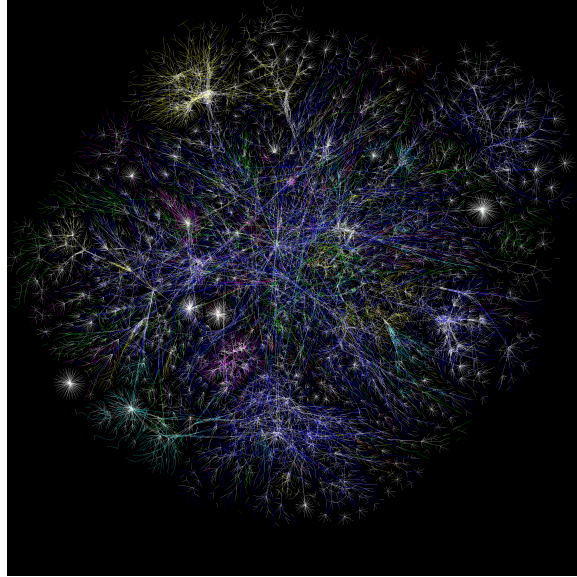


Figure 12: The Internet



Figure 13: Endre Szeméredi

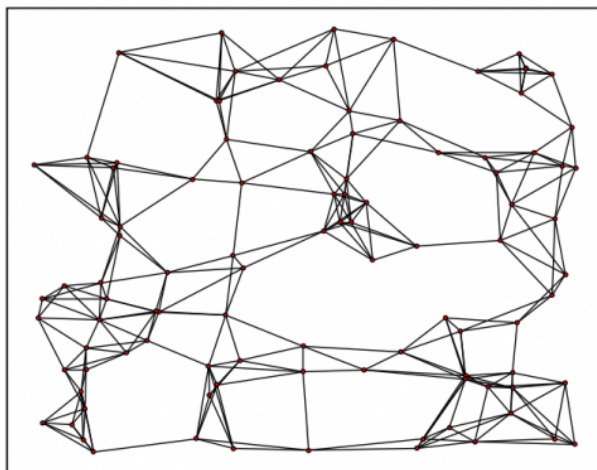


Figure 14: A network

of a complex network like the Internet. Now here's the question: How could I draw the Internet? What would be a good picture of Internet? Clearly, the number of pages on the Internet in the world is so big that it's impossible to draw it here. There are more points on Internet than pixels on the screen. So there's no way of drawing a picture of the actual Internet. How can I draw a fairly accurate picture of large networks? This is what Szemerédi's theorem tells you. It is possible to do something, and that's what I want to express.

Here's a network (Figure 15). You have dots and you have links between them, and the idea is that we want to group vertices or dots into several groups. We want to replace the complicated picture having many dots by a much simpler picture with fewer dots. Instead of having maybe twenty dots, you'll collapse these into only five dot-groups, A_1 , A_2 , A_3 , A_4 , A_5 , which you'll think of as new dots. Now let me give a definition and then state the theorem, since this is my job. My job is to state and then prove theorems.

So here's a similar picture (Figure 16). You have two sets, A and B . Inside A and B you have subsets, X_1 and Y_1 , where X_1 is subset of A , and Y_1 is subset of B .

And then we define some numbers, the first number is called the *density*.

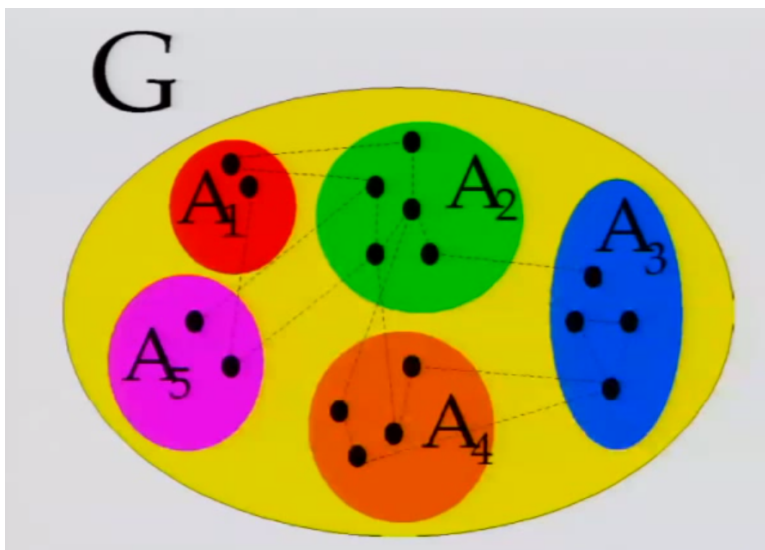


Figure 15: Simplifying a network

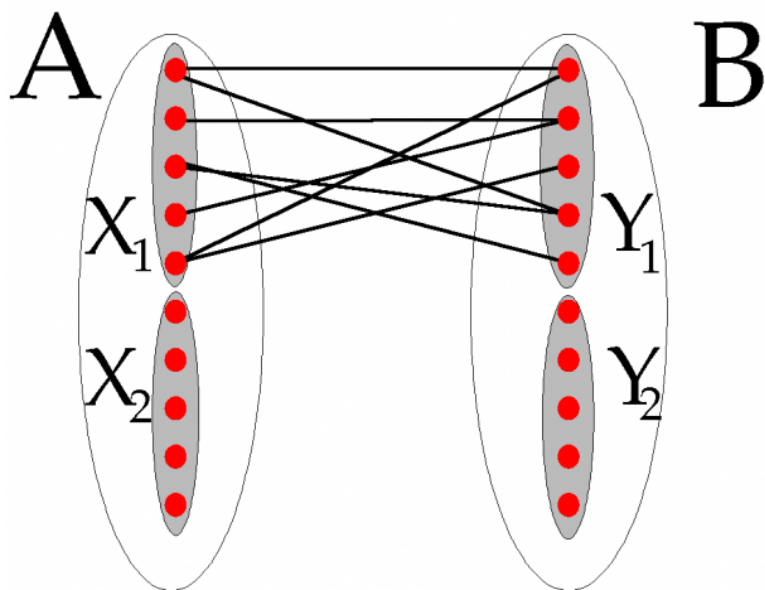


Figure 16: Defining ϵ -regularity

To calculate the density $d(X, Y)$ of X and Y , you count how many edges go from X to Y ; that is, the total number $e(X, Y)$ of connections going from some point of X to some point of Y . You then divide this number by the product of the number $|X|$ of points in X and the number $|Y|$ of points in Y .

$$d(X, Y) = \frac{e(X, Y)}{|X||Y|}$$

The density tells you the probability of connecting two points in X, Y . You say that two sets A, B are ϵ -regular if for every subset $X \subset A$ and $Y \subset B$, the density $d(X, Y)$ and the density $d(A, B)$ agree up to a small number $\epsilon > 0$. Here is the formal definition: A pair of sets (A, B) is ϵ -regular if for every $X \subset A$ with $|X| > \epsilon|A|$ and every $Y \subset B$ with $|Y| > \epsilon|B|$, we have

$$|d(X, Y) - d(A, B)| < \epsilon.$$

In other words, A and B are ϵ -regular if any part of A and any part of B are connected basically in the same way. Now I can state the theorem of Szemerédi, which says that every graph or network, for a given ϵ , can be approximated by a smaller graph, with a number of points independent of the original size of the graph, but only dependent on ϵ .

Endre Szemerédi regularity theorem. *For every $\epsilon > 0$, there are positive integers m and M such that every finite graph can be partitioned in n parts A_i in such a way that*

- $m \leq n \leq M$
- All A_i have approximately the same size: $(1-\epsilon)|A_i| \leq |A_j| \leq (1+\epsilon)|A_i|$
- Among the n^2 pairs (A_i, A_j) at least $(1-\epsilon)n^2$ are ϵ -regular.

This means that, whatever the size of the original network, you can approximate it by a small graph which gives you almost all the information you want about the connectivity inside your original network.

Let me end by showing you one example. This is a famous picture, an old picture, of the Internet (Figure 17). Of course, it's a very naïve image of Internet. It tells you that you can, roughly speaking, decompose the Internet into several parts. You have the “SCC”, that means the strongly connected core. It's about one-third of the total internet world. This is the part in which everybody interacted with everybody, it's highly connected. Then you have “OUT”, which are the pages, where everybody goes, but nothing

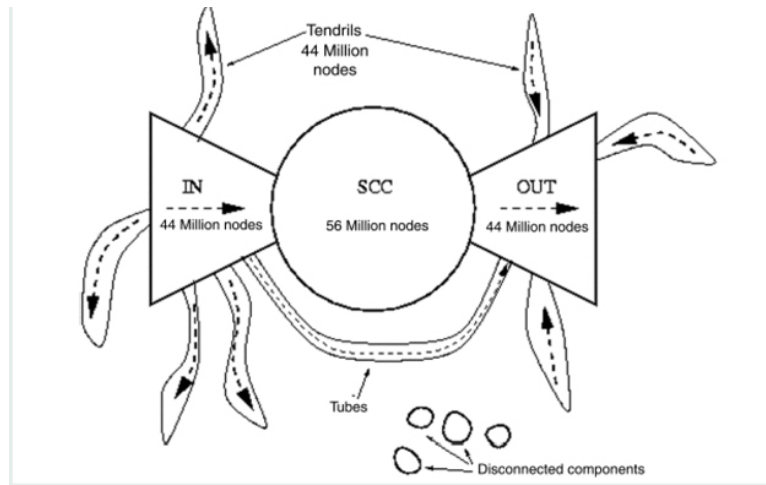


Figure 17: A small graph of the Internet

goes out of them. About the same size, you have “IN”, which consists of the pages that are not interesting to anybody, but which are interesting to everybody. Apart from these, there are some disconnected components, I don’t know what exactly those are, maybe the stamp collectors. The point is that this theorem of Szemerédi, in a word, explains that any network, even very big ones like the Internet network, can be described in such a way with a simple picture. It doesn’t tell you everything about the structure of the graph or network, but it tells you something about the global picture of it.

I wanted to mention this theorem to you primarily because it’s an example of a theorem for which the published proof is complicated, but nevertheless I understand it. For me it’s simple. I think I will never forget the proof because I understand it. And this is the exact opposite of the one-line by Jean-Pierre Serre, which was so short that it took me days to understand it. When you read the long proof of this theorem, once you get it, you will say “Well... I understand it, but why did they write such a long book on this?”

E.G
 UMPA UMR 5669 CNRS
 ENS Lyon
 Site Monod
 46 Allée d’Italie
 69364 Lyon
 France

etienne.ghys@ens-lyon.fr