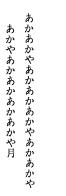# Mathematics as Metaphor

*Yuri Ivanovich Manin**

Steklov Mathematical Institute, 42 Vavilova, 117966 GSP-1, Moscow, USSR

あかあかや あかあかあかや あかあかや あかあかあかや あかあかや月

> *Ordre.* [...] *Je sais un peu ce que c'est et combien peu de gens l'entendent. Nulle science humaine ne le peut garder. Saint Thomas ne l'a pas gardé. La mathématique le garde, mais elle est inutile en sa profondeur.*
>
> Pascal, Pensées

Myōo (1173–1232)

Introduction

When H. Poincaré first published in 1902 his book *La Science et l'hypothèse*, it became a bestseller. The first chapter of this book was devoted to the nature of the mathematical reasoning. Poincaré discussed an old philosophical controversy whether the mathematical knowledge could be reduced to long chains of tautological transformations of some basic ("synthetic") truths or it contained something more. He argued that the creative power of mathematics was due to a free choice of the initial hypotheses-definitions which were later on constrained by the comparison of deductions with the observable world.

The society of our days seems to be much less interested in the philosophical subtleties than Poincaré's contemporaries. I do not want to say that science itself became less popular. Such books as S. Weinberg's *The first three minutes* and S.W. Hawking's *A brief history of time* are sold by hundreds of thousands and favorably reviewed in widely distributed newspapers. What has changed is the general mood. The paradoxality of the new physical theories is perceived less dramatically and more pragmatically. (We can note that the perception of visual arts evolved in much the same way: if the first exhibitions of Impressionists were a kind of spiritual revolution, each new wave of the post-war avant-garde immediately acquired family traits of academism).

---

* Delivered by Barry Mazur.

In this atmosphere, the heated discussions of the bygone days on the foundational crisis of mathematics and the nature of infinity seem almost irrelevant and certainly inappropriate. The audience responds much livelier to the opinions about school education or a new generation of computers.

This is why I have decided to present at this section an unpretentious essay in which our science is considered as a specialized dialect of the natural language, and its functioning as a special case of speech. This implies certain suggestions about the high school and University training.

Metaphorism

The word "metaphor" is used here in a non-technical sense, which is best rendered by the following quotations from James P. Carse's book *Finite and Infinite Games*:

"Metaphor is the joining of like to unlike such that one can never become the other."
    "At its root all language has the character of metaphor, because no matter what it intends to do, it remains language, and remains absolutely unlike whatever it is about".
    "The unspeakability of nature is the very possibility of language".

Considering mathematics as a metaphor, I want to stress that the interpretation of the mathematical knowledge is a highly creative act. In a way, mathematics is a novel about Nature and Humankind. One cannot tell precisely what mathematics teaches us, in much the same way as one cannot tell what exactly we are taught by "War and Peace". The teaching itself is submerged in the act of re-thinking this teaching.

This opinion seemingly disagrees with the time-honored tradition of applied mathematics in scientific and technological calculations.

In fact, I want only to restore a certain balance between the technological and the humanitarian sides of mathematics.

Two Examples

Let me try to illustrate the metaphoric potential of mathematics by discussing two disjoint subjects: the Kolmogorov complexity and the "Dictator Theorem" due to K. Arrow.

i) Kolmogorov's complexity of a natural number N is the length of a shortest program P that can generate N, or the length of a shortest code of N. A reader should imagine a way of coding integers which is a partial recursive function $f(P)$ taking natural values. Kolmogorov's theorem states that among all such functions there exist the most economical ones in the following sense: if $C_f(N)$ is the minimal value of P such that $f(P) = N$, then $C_f(N) \leq \text{const} \cdot C_g(N)$, where const. depends only on f, g but not N.

Since P can be reconstructed from its binary notation, the length $K_f(N)$ of the shortest program generating N is bounded by $\log_2 C_f(N)$. This function, or rather

the class of all such functions defined up to a bounded summand, is the Kolmogorov complexity.

First of all, $K(N) \le \log N + \text{const}$. Of course, this conforms nicely with the historical successes of the positional notation systems which provided us with the number generating programs of logarithmic length. However, there are arbitrary large integers whose Kolmogorov's complexity is much smaller than the length of their notation, e.g., $K(10^N) \le K(N) + \text{const}$. In general, when we use large numbers at all, we seemingly use only those which have a relatively small Kolmogorov's complexity. Even decimal decompositions of π which are, probably, the longest well-defined numbers ever produced by mathematicians, are Kolmogorov simple, because $K([10^N \pi]) \le \log N + \text{const}$. In general, small Kolmogorov complexity = high degree of organization.

On the other hand, almost all integers N have the complexity close to $\log N$. For example, if $f(P) = N$ for an optimal f, then $K(P)$ is equivalent to $\log P$. Such integers have many remarkable properties which we usually connect with "randomness".

Second, Kolmogorov's complexity can be easily defined for discrete objects which are not numbers, for example, Russian or English texts. Therefore, "War and Peace" has a pretty well defined measure of its complexity; the indeterminacy is connected with the choice of an optimal coding and seems to be pretty small if one chooses one of the small number of reasonable codings.

From this viewpoint, is "War and Peace" a highly organized or an almost random combinatorial object?

Third, Kolmogorov's complexity is a non-computable function. More precisely, if f is optimal, there is no recursive function $G(N)$ which would differ from $C_f(N)$ by $\exp(O(1))$. One can only bound complexity by computable functions.

I feel that Kolmogorov's complexity is a notion that is very essential to keep in mind in any discussion of the nature of human knowledge.

As long as the content of our knowledge is expressed symbolically (verbally, digitally, ...) there are physical restrictions on the volume of information that can be kept and handled. We always rely upon various methods of information compressing. Kolmogorov complexity puts absolute restrictions on the efficiency of such a compression. When we speak, say, of physical laws, expressed by the equations of motion, we mean that a precise description of the behaviour of a physical system can be obtained by translating these laws into a computer program. But the complexity of laws we can discover and use is clearly bounded. Can we be sure that there are no laws of arbitrary high complexity, even governing the "elementary" systems?

At this point, our discussion becomes totally un-mathematical, and before a mathematically-minded audience I must stop here. But such is the fate of any metaphor.

ii) Arrow's Dictator Theorem was discovered around 1950. Mathematically, it is a combinatorial statement describing certain functions with values in binary relations. Intuitively, it is a formalized discussion of the problem of Social Choice. Suppose that a lawmaker has to establish a law which governs the processing of

individual wills of voters into a collective decision. If the problem is to choose one of the two alternatives, the standard solution is do it by the majority of votes. However, usually there are more than two alternatives (imagine the funds allocation problems), and voters may be asked to order them according to their preferences. What should be the algorithm extracting the collective preference from any set of individual preferences? Arrow considered algorithms satisfying some natural and democratic axioms (e.g., when everybody prefers A to B, the society prefers A to B). Nevertheless, he discovered that when there are more than two alternatives, the only way to achieve a solution is to nominate a member of the society ("the Dictator") and to equate his personal preference order to the social one. (Actually, this is one of the versions of Arrow's theorem discovered later. Also, it refers to the case of a finite society; in the infinite case, the social decisions can be made by ultrafilters, appropriately called "the ruling hierarchies".)

In a way, this theorem illustrates the content of Jean-Jacque Rousseau's idea of a Contrat Social.

The fundamental inconsistency of the image of the ideal democratic choice can be illustrated by the following story referring to three voters and three alternatives. It is the story of three knights errant at the cross-roads with a stone before them. The inscription on the stone prophesies only losses: who goes to the left will lose his sword; who goes to the right will lose his horse; who goes straight will lose his head. The knights dismount and start taking council. In a Russian version of this story, the knights have names and personalities: the youngest and ardent Alyosha Popovich, the eldest and wisest Dobrynya Nikitich, and the slow peasant Ilya Muromets. So Alyosha values sword more than horse, and horse more than his head; Dobrynya values most his head, then his sword, then his horse; and Ilya prefers his horse to head to sword.

A reader will note that the three individual preference orders constitute one and the same cyclic order on the set of alternatives. As a result, one can decide by majority the choice between any two of the alternatives, but the union of these decisions will be inconsistent: the democratic procedure cannot provide us with a well-ordered list. The knights sigh and delegate the decision-making power to Dobrynya.

Does the Arrow theorem tell us something that we did not know beforehand? Yes, I think it does if we are ready to discuss it seriously, that is, to look closely at the combinatorial proof, to imagine the possible real life content of various assumptions and elementary logical steps made on the way, in general, to enhance our imprecise imagination by the rigid logic of a mathematical reasoning. We can understand better, for example, some tricks of policy-making and some pitfalls the society can leap whole-heartedly into (like accepting without questioning a list of alternatives imposed by a ruling hierarchy, although precisely the compilation of this list can be the central issue of the social decision making).

At this stage, we come to the main topic of our discussion: what distinguishes a mathematical discourse from a natural language discourse, why the Pascalian "ordre" came to reign over our specialized symbolic activities, and is it truly so "useless in its profundity"?

Language and Mathematics

A very interesting chapter of the interaction between mathematics and humanities started about thirty years ago when the first serious attempts of automatic translation were made. These first attempts were a painful failure, at least so for many optimists who believed that in this domain, there are no fundamental obstacles, and it remains only to overcome technical difficulties connected with the sheer amount of information to be processed. In other words, they took for granted that the translation is in principle performed by a not very complex algorithm which only must be made explicit and then implemented as a computer program.

This assumption is a nice example of a mathematical metaphor (actually, a specialization of the general "computer metaphor" used in the brain sciences).

This metaphor proved to be extremely fruitful for the theoretical linguistics in general because it forced linguists to start describing vocabulary, semantics, accidence, and syntax of human languages with unprecedented degree of explicitness and completeness. Some totally new notions and tools were discovered thanks to this program.

However, the successes of the automatic translation itself were (and still are) scanty. It became clear that written human speech is an extremely bad input data for any algorithmic processing planned as translation or even as a logical deduction. (I add this proviso because there is nothing special in human speech considered as a material for, say, statistical studies).

This fact can be considered as a universal property of human languages, and it deserves some attention. One must first of all reject as too naive a usual explanation that the universe of meanings of a human language is too vast and poorly structured to admit a well organized metalanguage describing this world. The point is that even if we severely restrict this universe to the subset of arithmetic of small integer quantities, we shall still have to face the same difficulty. In fact, this difficulty was a decisive reason for the crystallization of the whole system of arithmetical notation and the basic algorithms of calculation, and later on of symbolic algebra. Even the vocabulary of elementary arithmetic in human language is basically archaic: the finite natural series of primitive societies "one, two, three, indefinitely many" is reproduced in the exponential scale in our "thousand, million, billion, zillion". The expressions for relatively small numbers like "1989" are actually names of the decimal notation and not of the numbers themselves.

The advantage of F. Viete's algebra over the semi-verbal algebra of Diophantus was due not to the fact that it could express new meanings but to the incomparably greater susceptibility to the algorithmic processing ("identical transformations" of our high school algebra).

The rupture of the intuitive and emotional ties between a text and its producer/ user so characteristic of the language of science was compensated by the new computational automatisms. In their (albeit restricted) domain they proved to be infinitely more efficient that the traditional Platonian and Aristotelian culture of everyday language discourse. Why then our scientific papers are still written as a disorganized mixture of words and formulas? Partly because we still need those emotional ties; partly becaue some meanings (like human values) are best rendered

in human language. But even as a medium of scientific speech, human language has some inherent advantages: appealing to the spatial and qualitative imagination, it helps to understand "structurally stable" properties like the number of free parameters (dimension), existence of extrema, symmetries. To put it bluntly, it makes possible the metaphorical use of science.

Metaphor and Proof

The views professed here can be considered in relation to the high school and graduate curriculum.

The general mathematical education of the first half of this century was application oriented. It provided the basic minimum for the practical life problems and a smooth transition to the study of engineering and scientific calculations at the college level. The break of this curriculum with the activity of professional mathematicians became more and more pronounced. As is well known, this brought the reaction in the form of NewMath in USA and similar programs in other countries. These programs introduced into the high school mathematics the notions and principles borrowed from professionals: set theory, axiomatic presentation of proofs, strict culture of definitions.

NewMath became widely accepted but its expansion was accompanied by the protesting voices which in the 70s and 80s merged into a loud chorus. The critics disagreed with the basic arguments of the NewMath proponents. Leaving aside the objections based upon the data from cognitive sciences and learning psychology I shall only recall those concerning the general evaluation of the role of the proof in mathematics.

The one pole is represented by the well known statement due to Nicolas Bourbaki: "Dépuis les Grécs, qui dit Mathématiques, dit démonstration". According to this perception, the rigorous proof was made a matter of principle in the NewMath programs. It was argued that: a) a proof helps to understand a mathematical fact; b) a rigorous proof is the most essential component of the modern professional mathematics; c) mathematics possesses the universally recognized criteria of rigour.

These views were extensively criticised, e.g. by Gila Hanna in the book "Rigorous Proof in Mathematics Education", OISE Press, Ontario 1983. In particular, Gila Hanna pointed out that the mathematicians are far from unanimous in accepting the criteria of rigour (referring to quarrels between logicists, formalists and intuitionists), and that working mathematicians constantly break all rules in the book.

In my opinion, this is irrelevant.

What is relevant, is the imbalance between various basic values which is produced by the emphasis on proof. Proof itself is a derivate of the notion of "truth". There are a lot of values besides truth, among them "activities", "beauty" and "understanding", which are essential in the high school teaching and later. Neglecting precisely these values, a teacher (or a university professor) tragically fails. Unfortunately, this also is not universally recognized. A sociological analysis of the controversies around the Catastroph Theory of René Thom shows that exactly the

shift of orientation from the formal truth to understanding provoked such a sharp criticism. But of course, the Catastroph Theory is one of the developed mathematical metaphors and should only be judged as such.

Pedagogically, a proof is just one of the genres of a mathematical text. There are many different genres: a calculation, a commented sketch, a computer program, a description of an algorithmic language, or such a neglected kind as a discussion of the connections between a formal definition and intuitive notions. Every genre has its own laws, in particular, laws of rigour, which only are not codified because they were not payed a special attention.

A central problem of a teacher is to demonstrate at the restricted area of his or her course the variety of types of mathematical activities and underlying value orientations. Of course, this variety is hierarchically organized. The goals may vary from achieving an elementary arithmetical and logical literacy to programming skills, and from the simplest everyday problems to the principles of modern scientific thinking. In the spectrum of these goals, the emphasis on the norms of "rigorous proof" can safely occupy a peripheral position.

But having said all this, I must stress that my argumentation by no means undermines the ideal of a rigorous mathematical reasoning. This ideal is a fundamental constituting principle of mathematics, and in this sense Bourbaki is certainly right. Having no external object of study, being based on a consensus of a restricted circle of devotees, the mathematics could not develop without the permanent control of rigid rules of game. Applicability of mathematics in the strict sense of this word (like its indispensability in the Apollo project) is due to our ability to control series of symbolic manipulations of fantastic length.

The existence of this ideal is far more essential than its unattainability. The freedom of mathematics (G. Cantor) can only develop in the limits of iron necessity. The hardware of modern computers is an incarnation of this necessity.

Metaphor helps a human being to breathe in this rarefied atmosphere of Gods.