

The Mathematician *

John von Neumann

A discussion of the nature of intellectual work is a difficult task in any field, even in fields which are not so far removed from the central area of our common human intellectual effort as mathematics still is. A discussion of the nature of any intellectual effort is difficult per se — at any rate, more difficult than the mere exercise of that particular intellectual effort. It is harder to understand the mechanism of an airplane, and the theories of the forces which lift and which propel it, than merely to ride in it, to be elevated and transported by it — or even to steer it. It is exceptional that one should be able to acquire the understanding of a process without having previously acquired a deep familiarity with running it, with using it, before one has assimilated it in an instinctive and empirical way.

Thus any discussion of the nature of intellectual effort in any field is difficult, unless it presupposes an easy, routine familiarity with that field. In mathematics this limitation becomes very severe, if the discussion is to be kept on a non-mathematical plane. The discussion will then necessarily show some very bad features; points which are made can never be properly documented, and a certain over-all superficiality of the discussion becomes unavoidable.

I am very much aware of these shortcomings in what I am going to say, and I apologize in advance. Besides, the views which I am going to express are probably not wholly shared by many other mathematicians — you will get one man's not-too-well systematized impressions and interpretations — and I can give you only very little help in deciding how much they are to the point.

In spite of all these hedges, however, I must admit that it is an interesting and challenging task to make the attempt and to talk to you about the nature of intellectual effort in mathematics. I only hope that I will not fail too badly.

The most vitally characteristic fact about mathematics is, in my opinion, its quite peculiar relationship to the natural sciences, or, more generally, to any science which interprets experience on a higher than purely descriptive level.

Most people, mathematicians and others, will agree that mathematics is not an empirical science, or at least that it is practiced in a manner which differs in several deci-

* Published in *Works of the Mind*, by Robert B. Heywood, editor, vol. 1, no. 1, 180–196, University of Chicago Press, Chicago, 1947.

sive respects from the techniques of the empirical sciences. And, yet, its development is very closely linked with the natural sciences. One of its main branches, geometry, actually started as a natural, empirical science. Some of the best inspirations of modern mathematics (I believe, the best ones) clearly originated in the natural sciences. The methods of mathematics pervade and dominate the “theoretical” divisions of the natural sciences. In modern empirical sciences it has become more and more a major criterion of success whether they have become accessible to the mathematical method or to the near-mathematical methods of physics. Indeed, throughout the natural sciences an unbroken chain of successive pseudomorphoses, all of them pressing toward mathematics, and almost identified with the idea of scientific progress, has become more and more evident. Biology becomes increasingly pervaded by chemistry and physics, chemistry by experimental and theoretical physics, and physics by very mathematical forms of theoretical physics.

There is a quite peculiar duplicity in the nature of mathematics. One has to realize this duplicity, to accept it, and to assimilate it into one’s thinking on the subject. This double face is the face of mathematics, and I do not believe that any simplified, unitarian view of the thing is possible without sacrificing the essence.

I will therefore not attempt to present you with a unitarian version. I will attempt to describe, as best I can, the multiple phenomenon which is mathematics.

It is undeniable that some of the best inspirations in mathematics — in those parts of it which are as pure mathematics as one can imagine — have come from the natural sciences. We will mention the two most monumental facts.

The first example is, as it should be, geometry. Geometry was the major part of ancient mathematics. It is, with several of its ramifications, still one of the main divisions of modern mathematics. There can be no doubt that its origin in antiquity was empirical and that it began as a discipline not unlike theoretical physics today. Apart from all other evidence, the very name “geometry” indicates this. Euclid’s postulational treatment represents a great step away from empiricism, but it is not at all simple to defend the position that this was the decisive and final step, producing an absolute separation. That Euclid’s axiomatization does at some minor points not meet the modern requirements of absolute axiomatic rigour is of lesser importance in this respect. What is more essential, is this: other disciplines, which are undoubtedly empirical, like mechanics and thermodynamics, are usually presented in a more or less postulational treatment, which in the presentation of some authors is hardly distinguishable from Euclid’s procedure. The classic of theoretical physics in our time, Newton’s *Principia*, was, in literary form as well as in the essence of some of its most critical parts, very much like Euclid. Of course in all these instances there is behind the postulational presentation the physical insight backing the postulates and the experimental verification supporting the theorems. But one might well argue that a similar interpretation of Euclid is possible, especially from the viewpoint of antiquity, before geometry had acquired its

present bimillennial stability and authority — an authority which the modern edifice of theoretical physics is clearly lacking.

Furthermore, while the de-empirization of geometry has gradually progressed since Euclid, it never became quite complete, not even in modern times. The discussion of non-Euclidean geometry offers a good illustration of this. It also offers an illustration of the ambivalence of mathematical thought. Since most of the discussion took place on a highly abstract plane, it dealt with the purely logical problem whether the “fifth postulate” of Euclid was a consequence of the others or not; and the formal conflict was terminated by F Klein’s purely mathematical example, which showed how a piece of a Euclidean plane could be made non-Euclidean by formally redefining certain basic concepts. And yet the empirical stimulus was there from start to finish. The prime reason, why, of all Euclid’s postulates, the fifth was questioned, was clearly the unempirical character of the concept of the entire infinite plane which intervenes there, and there only. The idea that in at least one significant sense—and in spite of all mathematico-logical analyses—the decision for or against Euclid may have to be empirical, was certainly present in the mind of the greatest mathematician, Gauss. And after Bolyai, Lobachevsky, Riemann, and Klein had obtained more abstracto, what we today consider the formal resolution of the original controversy, empirics — or rather physics — nevertheless, had the final say. The discovery of general relativity forced a revision of our views on the relationship of geometry in an entirely new setting and with a quite new distribution of the purely mathematical emphases, too. Finally, one more touch to complete the picture of contrast. This last development took place in the same generation which saw the complete de-empirization and abstraction of Euclid’s axiomatic method in the hands of the modern axiomatic-logical mathematicians. And these two seemingly conflicting attitudes are perfectly compatible in one mathematical mind; thus Hilbert made important contributions to both axiomatic geometry and to general relativity.

The second example is calculus — or rather all of analysis, which sprang from it. The calculus was the first achievement of modern mathematics, and it is difficult to overestimate its importance. I think it defines more unequivocally than anything else the inception of modern mathematics, and the system of mathematical analysis, which is its logical development, still constitutes the greatest technical advance in exact thinking.

The origins of calculus are clearly empirical. Kepler’s first attempts at integration were formulated as “dolichometry” — measurement of kegs — that is, volumetry for bodies with curved surfaces. This is geometry, but post-Euclidean, and, at the epoch in question, non-axiomatic, empirical geometry. Of this, Kepler was fully aware. The main effort and the main discoveries, those of Newton and Leibniz, were of an explicitly physical origin. Newton invented the calculus “of fluxions” essentially for the purposes of mechanics — in fact, the two disciplines, calculus and mechanics, were developed by him more or less together. The first formulations of the calculus were not even mathematically rigorous. An inexact, semi-physical formulation was the only one available

for over a hundred and fifty years after Newton! And yet, some of the most important advances of analysis took place during this period, against this inexact, mathematically inadequate background! Some of the leading mathematical spirits of the period were clearly not rigorous, like Euler; but others, in the main, were, like Gauss or Jacobi. The development was as confused and ambiguous as can be, and its relation to empiricism was certainly not according to our present (or Euclid's) ideas of abstraction and rigour. Yet no mathematician would want to exclude it from the fold-that period produced mathematics as first class as ever existed! And even after the reign of rigour was essentially re-established with Cauchy, a very peculiar relapse into semi-physical methods took place with Riemann. Riemann's scientific personality itself is a most illuminating example of the double nature of mathematics, as is the controversy of Riemann and Weierstrass, but it would take me too far into technical matters if I went into specific details. Since Weierstrass, analysis seems to have become completely abstract, rigorous, and unempirical. But even this is not unqualifiedly true. The controversy about the "foundations" of mathematics and logics, which took place during the last two generations, dispelled many illusions on this score.

This brings me to the third example which is relevant for the diagnosis. This example, however, deals with the relationship of mathematics with philosophy or epistemology rather than with the natural sciences. It illustrates in a very striking fashion that the very concept of "absolute" mathematical rigour is not immutable. The variability of the concept of rigour shows that something else besides mathematical abstraction must enter into the makeup of mathematics. In analyzing the controversy about the "foundations," I have not been able to convince myself that the verdict must be in favour of the empirical nature of this extra component. The case in favour of such an interpretation is quite strong, at least in some phases of the discussion. But I do not consider it absolutely cogent. Two things, however, are clear. First, that something nonmathematical, somehow connected with the empirical sciences or with philosophy or both, does enter essentially-and its non-empirical character could only be maintained if one assumed that philosophy (or more specifically epistemology) can exist independently of experience. (And this assumption is only necessary but not in itself sufficient). Second, that the empirical origin of mathematics is strongly supported by instances like our two earlier examples (geometry and calculus), irrespective of what the best interpretation of the controversy about the "foundations" may be.

In analyzing the variability of the concept of mathematical rigour, I wish to lay the main stress on the "foundations" controversy, as mentioned above. I would, however, like to consider first briefly a secondary aspect of the matter. This aspect also strengthens my argument, but I do consider it as secondary, because it is probably less conclusive than the analysis of the "foundations" controversy. I am referring to the changes of mathematical "style." It is well known that the style in which mathematical proofs are written has undergone considerable fluctuations. It is better to talk of fluctuations than

of a trend because in some respects the difference between the present and certain authors of the eighteenth or of the nineteenth centuries is greater than between the present and Euclid. On the other hand, in other respects there has been remarkable constancy. In fields in which differences are present, they are mainly differences in presentation, which can be eliminated without bringing in any new ideas. However, in many cases these differences are so wide that one begins to doubt whether authors who “present their cases” in such divergent ways can have been separated by differences in style, taste, and education only—whether they can really have had the same ideas as to what constitutes mathematical rigour. Finally, in the extreme cases (e.g., in much of the work of the late-eighteenth-century analysis, referred to above), the differences are essential and can be remedied, if at all, only with the help of new and profound theories, which it took up to a hundred years to develop. Some of the mathematicians who worked in such, to us, unrigorous ways (or some of their contemporaries, who criticized them) were well aware of their lack of rigour. Or to be more objective: Their own desires as to what mathematical procedure should be were more in conformity with our present views than their actions. But others — the greatest virtuoso of the period, for example, Euler — seem to have acted in perfect good faith and to have been quite satisfied with their own standards.

However, I do not want to press this matter further. I will turn instead to a perfectly clear-cut case, the controversy about the “foundations of mathematics.” In the late nineteenth and the early twentieth centuries a new branch of abstract mathematics, G Cantor’s theory of sets, led into difficulties. That is, certain reasonings led to contradictions; and, while these reasonings were not in the central and “useful” part of set theory, and always easy to spot by certain formal criteria, it was nevertheless not clear why they should be deemed less set-theoretical than the “successful” parts of the theory. Aside from the *ex post* insight that they actually led into disaster, it was not clear what a priori motivation, what consistent philosophy of the situation, would permit one to segregate them from those parts of set theory which one wanted to save. A closer study of the merits of the case, undertaken mainly by Russell and Weyl, and concluded by Brouwer, showed that the way in which not only set theory but also most of modern mathematics used the concepts of “general validity” and of “existence” was philosophically objectionable. A system of mathematics which was free of these undesirable traits, “intuitionism,” was developed by Brouwer. In this system the difficulties and contradiction of set theory did not arise. However, a good fifty per cent of modern mathematics, in its most vital — and up to then unquestioned — parts, especially in analysis, were also affected by this “purge”: they either became invalid or had to be justified by very complicated subsidiary considerations. And in this latter process one usually lost appreciably in generality of validity and elegance of deduction. Nevertheless, Brouwer and Weyl considered it necessary that the concept of mathematical rigour be revised according to these ideas.

It is difficult to overestimate the significance of these events. In the third decade of the twentieth century two mathematicians—both of them of the first magnitude, and as deeply and fully conscious of what mathematics is, or is for, or is about, as anybody could be—actually proposed that the concept of mathematical rigour, of what constitutes an exact proof, should be changed! The developments which followed are equally worth noting.

1. Only very few mathematicians were willing to accept the new, exigent standards for their own daily use. Very many, however, admitted that Weyl and Brouwer were *prima facie* right, but they themselves continued to trespass, that is, to do their own mathematics in the old, “easy” fashion—probably in the hope that somebody else, at some other time, might find the answer to the intuitionistic critique and thereby justify them *a posteriori*.

2. Hilbert came forward with the following ingenious idea to justify “classical” (i.e., pre-intuitionistic) mathematics: Even in the intuitionistic system it is possible to give a rigorous account of how classical mathematics operate, that is, one can describe how the classical system works, although one cannot justify its workings. It might therefore be possible to demonstrate intuitionistically that classical procedures can never lead into contradictions—into conflicts with each other. It was clear that such a proof would be very difficult, but there were certain indications how it might be attempted. Had this scheme worked, it would have provided a most remarkable justification of classical mathematics on the basis of the opposing intuitionistic system itself! At least, this interpretation would have been legitimate in a system of the philosophy of mathematics which most mathematicians were willing to accept.

3. After about a decade of attempts to carry out this program, Gödel produced a most remarkable result. This result cannot be stated absolutely precisely without several clauses and caveats which are too technical to be formulated here. Its essential import, however, was this: If a system of mathematics does not lead into contradiction, then this fact cannot be demonstrated with the procedures of that system. Gödel’s proof satisfied the strictest criterion of mathematical rigour — the intuitionistic one. Its influence on Hilbert’s program is somewhat controversial, for reasons which again are too technical for this occasion. My personal opinion, which is shared by many others, is, that Gödel has shown that Hilbert’s program is essentially hopeless.

4. The main hope of a justification of classical mathematics — in the sense of Hilbert or of Brouwer and Weyl — being gone, most mathematicians decided to use that system anyway. After all, classical mathematics was producing results which were both elegant and useful, and, even though one could never again be absolutely certain of its reliability, it stood on at least as sound a foundation as, for example, the existence of the electron. Hence, if one was willing to accept the sciences, one might as well accept the classical system of mathematics. Such views turned out to be acceptable even to some of the original protagonists of the intuitionistic system. At present the controversy about

the “foundations” is certainly not closed, but it seems most unlikely that the classical system should be abandoned by any but a small minority.

I have told the story of this controversy in such, detail, because I think that it constitutes the best caution against taking the immovable rigour of mathematics too much for granted. This happened in our own lifetime, and I know myself how humiliatingly easily my own views regarding the absolute mathematical truth changed during this episode, and how they changed three times in succession!

I hope that the above three examples illustrate one-half of my thesis sufficiently well-that much of the best mathematical inspiration comes from experience and that it is hardly possible to believe in the existence of an absolute, immutable concept of mathematical rigour, dissociated from all human experience. I am trying to take a very low-brow attitude on this matter. Whatever philosophical or epistemological preferences anyone may have in this respect, the mathematical fraternities’ actual experiences with its subject give little support to the assumption of the existence of an a priori concept of mathematical rigour. However, my thesis also has a second half, and I am going to turn to this part now.

It is very hard for any mathematician to believe that mathematics is a purely empirical science or that all mathematical ideas originate in empirical subjects. Let me consider the second half of the statement first. There are various important parts of modern mathematics in which the empirical origin is untraceable, or, if traceable, so remote that it is clear that the subject has undergone a complete metamorphosis since it was cut off from its empirical roots. The symbolism of algebra was invented for domestic, mathematical use, but it may be reasonably asserted that it had strong empirical ties. However, modern, “abstract” algebra has more and more developed into directions which have even fewer empirical connections. The same may be said about topology. And in all these fields the mathematician’s subjective criterion of success, of the worthwhileness of his effort, is very much self-contained and aesthetical and free (or nearly free) of empirical connections. (I will say more about this further on.) In set theory this is still clearer. The “power” and the “ordering” of an infinite set may be the generalizations of finite numerical concepts, but in their infinite form (especially “power”) they have hardly any relation to this world. If I did not wish to avoid technicalities, I could document this with numerous set theoretical examples-the problem of the “axiom of choice,” the “comparability” of infinite “powers,” the “continuum problem,” etc. The same remarks apply to much of real function theory and real point-set theory. Two strange examples are given by differential geometry and by group theory: they were certainly conceived as abstract, non-applied disciplines and almost always cultivated in this spirit. After a decade in one case, and a century in the other, they turned out to be very useful in physics. And they are still mostly pursued in the indicated, abstract, non-applied spirit.

The examples for all these conditions and their various combinations could be multiplied, but I prefer to turn instead to the first point I indicated above: Is mathematics an empirical science? Or, more precisely: Is mathematics actually practiced in the way in which an empirical science is practiced? Or, more generally: What is the mathematician's normal relationship to his subject? What are his criteria of success, of desirability? What influences, what considerations, control and direct his effort?

Let us see, then, in what respects the way in which the mathematician normally works differs from the mode of work in the natural sciences. The difference between these, on one hand, and mathematics, on the other, goes on, clearly increasing as one passes from the theoretical disciplines to the experimental ones and then from the experimental disciplines to the descriptive ones. Let us therefore compare mathematics with the category which lies closest to it — the theoretical disciplines. And let us pick there the one which lies closest to mathematics. I hope that you will not judge me too harshly if I fail to control the mathematical hybris and add: because it is most highly developed among all theoretical sciences—that is, theoretical physics. Mathematics and theoretical physics have actually a good deal in common. As I have pointed out before, Euclid's system of geometry was the prototype of the axiomatic presentation of classical mechanics, and similar treatments dominate phenomenological thermodynamics as well as certain phases of Maxwell's system of electrodynamics and also of special relativity. Furthermore, the attitude that theoretical physics does not explain phenomena, but only classifies and correlates, is today accepted by most theoretical physicists. This means that the criterion of success for such a theory is simply whether it can, by a simple and elegant classifying and correlating scheme, cover very many phenomena, which without this scheme would seem complicated and heterogeneous, and whether the scheme even covers phenomena which were not considered or even not known at the time when the scheme was evolved. (These two latter statements express, of course, the unifying and the predicting power of a theory.) Now this criterion, as set forth here, is clearly to a great extent of an aesthetical nature. For this reason it is very closely akin to the mathematical criteria of success, which, as you shall see, are almost entirely aesthetical. Thus we are now comparing mathematics with the empirical science that lies closest to it and with which it has, as I hope I have shown, much in common — with theoretical physics. The differences in the actual *modus procedendi* are nevertheless great and basic. The aims of theoretical physics are in the main given from the “outside,” in most cases by the needs of experimental physics. They almost always originate in the need of resolving a difficulty; the predictive and unifying achievements usually come afterward. If we may be permitted a simile, the advances (predictions and unifications) come during the pursuit, which is necessarily preceded by a battle against some pre-existing difficulty (usually an apparent contradiction within the existing system). Part of the theoretical physicists' work is a search for such obstructions, which promise a possibility for a “break-through.” As I mentioned, these difficulties originate usually

in experimentation, but sometimes they are contradictions between various parts of the accepted body of theory itself. Examples are, of course, numerous.

Michelson's experiment leading to special relativity, the difficulties of certain ionization potentials and of certain spectroscopic structures leading to quantum mechanics exemplify the first case; the conflict between special relativity and Newtonian gravitational theory leading to general relativity exemplifies the second, rarer, case. At any rate, the problems of theoretical physics are objectively given; and, while the criteria which govern the exploitation of a success are, as I indicated earlier, mainly aesthetical, yet the portion of the problem, and that which I called above the original "break-through," are hard, objective facts. Accordingly, the subject of theoretical physics was at almost all times enormously concentrated; at almost all times most of the effort of all theoretical physicists was concentrated on no more than one or two very sharply circumscribed fields—quantum theory in the 1920's and early 1930's and elementary particles and structure of nuclei since the mid-1930's are examples.

The situation in mathematics is entirely different. Mathematics falls into a great number of subdivisions, differing from one another widely in character, style, aims, and influence. It shows the very opposite of the extreme concentration of theoretical physics. A good theoretical physicist may today still have a working knowledge of more than half of his subject. I doubt that any mathematician now living has much of a relationship to more than a quarter. "Objectively" given, "important" problems may arise after a subdivision of mathematics has evolved relatively far and if it has bogged down seriously before a difficulty. But even then the mathematician is essentially free to take it or leave it and turn to something else, while an "important" problem in theoretical physics is usually a conflict, a contradiction, which "must" be resolved. The mathematician has a wide variety of fields to which he may turn, and he enjoys a very considerable freedom in what he does with them. To come to the decisive point: I think that it is correct to say that his criteria of selection, and also those of success, are mainly aesthetical. I realize that this assertion is controversial and that it is impossible to "prove" it, or indeed to go very far in substantiating it, without analyzing numerous specific, technical instances. This would again require a highly technical type of discussion, for which this is not the proper occasion. Suffice it to say that the aesthetical character is even more prominent than in the instance I mentioned above in the case of theoretical physics. One expects a mathematical theorem or a mathematical theory not only to describe and to classify in a simple and elegant way numerous and a priori disparate special cases. One also expects "elegance" in its "architectural," structural makeup. Ease in stating the problem, great difficulty in getting hold of it and in all attempts at approaching it, then again some very surprising twist by which the approach, or some part of the approach, becomes easy, etc. Also, if the deductions are lengthy or complicated, there should be some simple general principle involved, which "explains" the complications and detours, reduces the apparent arbitrariness to a few simple guiding motivations, etc.

These criteria are clearly those of any creative art, and the existence of some underlying empirical, worldly motif in the background — often in a very remote background — overgrown by aestheticizing developments and followed into a multitude of labyrinthine variants — all this is much more akin to the atmosphere of art pure and simple than to that of the empirical sciences.

You will note that I have not even mentioned a comparison of mathematics with the experimental or with the descriptive sciences. Here the differences of method and of the general atmosphere are too obvious.

I think that it is a relatively good approximation to truth — which is much too complicated to allow anything but approximations—that mathematical ideas originate in empirics, although the genealogy is sometimes long and obscure. But, once they are so conceived, the subject begins to live a peculiar life of its own and is better compared to a creative one, governed by almost entirely aesthetic motivations, than to anything else and, in particular, to an empirical science. There is, however, a further point which, I believe, needs stressing. As a mathematical discipline travels far from its empirical source, or still more, if it is a second and third generation only indirectly inspired by ideas coming from “reality” it is beset with very grave dangers. It becomes more and more purely aestheticizing, more and more purely *l’art pour l’art*. This need not be bad, if the field is surrounded by correlated subjects, which still have closer empirical connections, or if the discipline is under the influence of men with an exceptionally well-developed taste. But there is a grave danger that the subject will develop along the line of least resistance, that the stream, so far from its source, will separate into a multitude of insignificant branches, and that the discipline will become a disorganized mass of details and complexities. In other words, at a great distance from its empirical source, or after much “abstract” inbreeding, a mathematical subject is in danger of degeneration. At the inception the style is usually classical; when it shows signs of becoming baroque, then the danger signal is up. It would be easy to give examples, to trace specific evolutions into the baroque and the very high baroque, but this, again, would be too technical.

In any event, whenever this stage is reached, the only remedy seems to me to be the rejuvenating return to the source: the re-injection of more or less directly empirical ideas. I am convinced that this was a necessary condition to conserve the freshness and the vitality of the subject and that this will remain equally true in the future.