Interrelations between Mathematics and Physics

Yu. I. Manin*

Abstract

After briefly describing the mathematical structure of modern physics, this paper analyzes the divergence between the development of physics and of mathematics in the first half of the 20th century, with emphasis on the role in each discipline of rigorous definitions and proofs, of algebraic calculations and of intuitive ideas.

Résumé

Après avoir décrit la structure mathématique de la physique moderne, cet article analyse la divergence entre mathématiques et physique dans la première moitié du XX^{e} siècle, en étudiant, pour chacune des disciplines, le rôle respectif des définitions et démonstrations rigoureuses, des calculs algébriques et des idées intuitives.

1. Foreword

I would like to start with an explicit description of the conceptual framework of this study.

To render it concisely, it is useful to look at the case of comparative linguistics. The history of a language is not a history of all, or even of "the most important," utterances (oral or written) in this language. Rather, it is a history of evolution *of the language as a system.* Hence we need a preliminary description of the system(s) whose genesis we are studying.

AMS 1991 *Mathematics Subject Classification*: 00A30, 01A60, 01A65

*Max–Planck–Institut für Mathematik, Bonn, Deutschland.

An application of this Saussurian scheme to the history of mathematics (which, incidentally, I do not consider to be a mere language) was probably particularly appealing to Jean Dieudonné who, as an active member of the Bourbaki group, participated in the creation of a systematic picture of modern mathematics.[1] In this talk I follow his example, on a much humbler scale. Needless to say that restrictions of time, space, and competence, force me to choose a thin chain of connected ideas and present them in a highly selective way.

Thus I refuse (somewhat reluctantly) to discuss the history with Rankean insistence on *wie es eigentlich gewesen ist.* One reason for this refusal is that the history of contemporary mathematics tends to degenerate into credit and priority assignments, lacking pathetically the dramatic appeal with which the history of struggles for real power is charged. A more personal and compelling motive is succinctly put by Joseph Brodsky in his autobiographical essay *Less Than One*: "The little I remember becomes even more diminished by being recollected in English."

A last word of warning and apology is due. Any system is, of course, a theoretical construct. As such, it is at best relative and culture dependent, at worst subjective. It is precisely in this function that it can serve as material for the history of mathematics of the 20th century.

2. Mathematical Physics as a System

2.1. Physics

Physics describes the external world, and in its domain of competence, does this in two complementary modes: classical and quantum.

In the *classical mode,* events occur to the matter and fields which reside and evolve in the space–time. Physical laws directly constrain observables. They are basically deterministic and expressed by the differential equations which (sometimes demonstrably, sometimes hypothetically) satisfy appropriate uniqueness and existence theorems.

A statistical submode of the classical mode of description deals with probabilities and averages which (sometimes demonstrably, sometimes presumably) can be deduced from an ideal deterministic description. The need for a statis-

[1] Jean Dieudonné, as I remember him, had a strong voice, strong hands, and strong opinions. In particular, he insisted on using tensor products and commutative diagrams instead of classical subscripts and superscripts in calculations involving tensors. I used to believe his judgement that this was a chalk–saving device, until one day I had to calculate with tensors myself. Then I found out that subscripts were much more economical.

tical treatment arises from two basic premises: too many degrees of freedom and/or instability. (Metaphorically speaking, instability means that each consecutive decimal digit is a new degree of freedom.)

A fundamental physical abstraction is that of *an isolated system* which evolves in oblivion of the rest of the world, and of *interaction* between potentially isolated systems, or one isolated system and the rest of the world.

In one of the most remarkable flights of fancy of classical physics, *space–time* itself appears as such an isolated system governed by Einstein's equations of general relativity (perhaps, with an energy–momentum tensor summarily responsible for everything which is not pure space–time).

In the *quantum mode* of theoretical description, the observable world is inherently probabilistic. Moreover, and more significantly, the basic laws — which are in a sense deterministic — govern an unobservable entity, the *probability amplitude*, which is a complex valued function on a quantum path space. Roughly speaking, the amplitude of a composite event is the product of the amplitudes of its constituents, whereas the amplitude of an event which is a sum of alternatives is the sum of the amplitudes of these alternatives.

The probability of an event is the modulus squared of its amplitude. Physical observables are the appropriate averages, even if one speaks about an elementary act of scattering of an individual particle. The observable wave behavior of, say, light is only an imperfect reflection of the inherent wave behavior of the amplitudes (wave functions) of an indeterminate number of photons described by the Fock space of the quantized electromagnetic field.

Partly as a result of historical development, many quantum models contain as an intermediate stage a classical model which is then quantized. The word "quantization" rather indiscriminately refers to a wide variety of procedures of which two of the most important are operator, or Hamiltonian, quantization, and *path integral* quantization. The first is more algebraic and usually has a firmer mathematical background. The second possesses an enormous heuristic and aesthetic potential. I haven chosen the latter for my more detailed subsequent discussion.

If I had included the first one, the picture of the divergence of Mathematics and Physics in the first half of this century sketched below in Sec. IV would appear less pronounced. Nevertheless, the main results of my analysis would survive.

One more subject matter deserving a separate historical and structural study is the duality between these two approaches. It started with classical mechanics, Lagrange, and Hamilton, and continued via Heisenberg–Schrödinger wave mechanics to the path integral/scattering matrix controversy. On the fringes of physics it contains such recent mathematical gems as

Virasoro algebra representations on the moduli spaces of curves.

2.2. Mathematics

If there is one most important notion of mathematical physics, it is that of *action* functional. It encompasses the classical ideas of energy and work, its density in a domain of space–time is the Lagrangian, and multiplied by $\sqrt{-1}$ and exponentiated, it furnishes the basic probability amplitude. Action is measured in absolute Planck units, and therefore can be thought of as a real number. More precisely, we will consider the following scheme of description central for both modes of physical description referred to above.

The modeling of a physical system starts with the specification of its kinematics. This includes a space $\mathcal{P}$ of virtual classical paths of the system and an action functional $S : \mathcal{P} \to \mathbb{R}$. For example, $\mathcal{P}$ may consist of parametrized curves in a classical phase space of a mechanical system, or of Riemannian metrics on a given smooth manifold (space–time), or of triples *(a connection on a given vector bundle, a metric on it, a section of it)* etc. The value of the action functional at a point $p \in \mathcal{P}$ is usually given in the form $\int_p L$, that is a volume form integrated over one of the spaces figuring in the description of p.

Classical equations of motion specify a subspace $\mathcal{P}_{cl} \subset \mathcal{P}$. This subset consists of the solutions of the variational equations $\delta(S) = 0$, *i. e.,* of the stationary points of the action functional.

If the classical description is the statistical one, then $\exp(-S)$ is the probability density.

In the quantum description, we choose physically motivated subsets $B \subset \mathcal{P}$, typically determined by boundary conditions, and define the average of an observable O in B by a path integral of the type

$$\langle O \rangle_B := \int_B O(p)\, e^{i\int_p L}\, Dp. \tag{2.1}$$

These are our main actors. In the following, I present some musings about the history of this picture as seen through the eyes of physicists and mathematicians.

I will be most interested in the idea of the integral and its final incarnation, in the form of the *path integral.*

3. The Integral

The notion of an integral is one of the central and recurring themes in the history of mathematics for the last two millennia. The ardent problem solving is periodically followed by the anxious definition seeking, only to be replaced by new non–rigorous but amazingly efficient heuristics leaving a logically–minded fundamentalist in each of us baffled.

Richard Feynman who created the hierogram (2.1) (still lacking a precise mathematical meaning exactly in those cases when it is most needed by physicists[2]) used to boast that (2.1) allowed the calculation of the anomalous magnetic momentum of the electron, which coincided with its experimental value up to ten digits:

"As of 1983, the theoretical number was 1.00115965246, with an uncertainty of about 20 in the last two digits; the experimental number was 1.00115965221, with an uncertainty of about 4 in the last digit. This accuracy is equivalent to measuring the distance from Los Angeles to New York, a distance of over 3000 miles, to within the width of a human hair." [Feynman 1988, p. 118]

This feat was recently matched by physical calculations (even called "predictions", cf. [Candelas *et al.* 1991]) of various interesting numbers in algebraic geometry, such as the number N_d of rational curves of degree d on a generic three–dimensional quintic (e. g. 70428 81649 78454 68611 34882 49750 for $d = 10$, a theoretical(?) number still unchecked in an experiment(?) involving a mathematical definition of N_d and a computer.) The ideology of path integration played an essential role in these calculations, leading to an interpretation of an instance of (2.1) as a sum over instantons in a sigma–model, which in this particular case are rational curves on a quintic.

The intuitive physical picture of an integral is *the quantity of something in a domain.* If the first calculations of this "something" are later interpreted as, say, the volume of a pyramid, one can hardly doubt that they were used for estimating the actual quantity of *stone* (and slaves' labor) needed for the building of an Egyptian pharaoh's tomb. Kepler's *Stereometria Doliorum* mentions *wine casks* in its title. The domain in question acquired a temporal dimension when *the length* of a path was calculated as an integral of velocity, and the notion of energy was gradually replaced by that of *action.* In the twentieth century, *topology* became one of the substances the quantity of which could be measured by integration of closed differential forms (De Rham theory of periods anticipated by Poincaré). *Probability* turned out to be another such substance, and Wiener's treatment of Brownian motion as a measure in a space of continuous paths paved the way both for Kolmogorov's axiomatic

[2]For a more positive view, see [Glimm and Jaffe 1981], a remarkable book which influenced the structure of this essay. On page 313 however the authors say: "... it is a theoretical puzzle whether a *theory* of electrodynamics exists in the sense of a mathematical framework ..."

treatment of probability and our present reluctant acceptance of Feynman's integral. (This is at least partially supported by the successes of constructive field theory and stochastic integration. However, the random surfaces inherent in string path integrals present considerable difficulties.)

Mathematically, any calculation (or definition) of an integral is based upon two physically intuitive principles: additivity with respect to domains and integrands, and a form of limiting procedure. There are at least two archetypal forms of passing to a limit.

One is represented by Cavalieri's indivisibles, Riemann sums, etc. It is connected with the topological structure of the domain of integration, specifically with the idea of boundary and thin layers of $(d+1)$–dimensional objects surrounding a d–dimensional object. The Stokes formula in all its modifications belongs to this circle of ideas, while the De Rham complex is its linear dual form.

Another form of limiting procedure is measure–theoretical rather than a topological one. There are basic domains filled with well measured quantities of the substance of interest (volume, action, probability ...). We try to approximate other distributions by using mosaic portraits of them and allowing the size of local discrepancies to tend to zero. However, locality is not topological anymore, and the image of boundary becomes useless or irrelevant. Instead, we have to deal with measurable sets which must only form an algebra with respect to intersections and unions. Infinite–dimensional constructions are usually of this type. The well known effect "volume in high dimensions tends to concentrate near the boundary" prevents using the image of indivisibles effectively. Even in finite dimensions, the boundary can fail to serve the role of Cavalieri's indivisible if it is very rough (fractal). The subtle measure theoretic studies of the beginning of this century had much to say about it.

There are *two* integrals in (2.1), of quite different nature. The action $S = \int_p L$ is usually a classical entity, L being a local Lagrangian. A beautiful recent idea due to a collaboration of physicists and mathematicians (E. Witten [1989] and M. F. Atiyah [1989] playing leading roles, A. S. Schwarz having supplied a crucial first example) consisted in considering those path integrals in which the action is a topological invariant of p. Locally this means that classical equations of motion $\delta(S) = 0$ are identically satisfied. An example of such an action functional is the Chern–Simons invariant defined on the space of connections on a vector bundle over a three–dimensional manifold. The quantum observables (whose choice and name was motivated by the theory of strong interactions) are Wilson loops: averaged traces of monodromy representations along closed curves in the base.

In this context, the algebraic properties of the path integral reflected in the additivity of $\int_p L$ and resulting "multiplicativity" of the whole of (2.1) become so strong that they can be used to define a sufficiently rigid mathematical structure of "Topological Quantum Field Theory" which can then be studied by precise mathematical means. This was done by G. Segal and M. F. Atiyah. See [Reshetikhin and Turaev 1991] and [Blanchet *et al.* 1995] for some recent mathematical developments in this area.

The history of the integral seen from our vantage point can be conceived in terms of a Toynbeean challenge/response scheme. Challenges come from physics broadly construed, including geometry. It can be convincingly argued that even Euclidean geometry is in fact just the kinematics of rigid bodies in the absence of a gravitational field (curved the space–time), and both the invention and the development of the first non–Euclidean geometries (of constant curvature) was inextricably connected with physics. Gauss wanted to know what was the *actual* geometry of interstellar space. Hilbert's return to axiomatics was a mathematical response to the challenge of the discovery of multiple *possible* geometries of the physical world.

4. The Schism

In this section of my talk I argue that the main event in the relationship between mathematics and physics in the first half of this century was their estrangement, after several centuries of close alliance.

The divergence started in the last two decades of the last century and was connected with the deepening understanding of two microworlds: a mathematical one embodied in the idea of the classical continuum of real numbers, and a physical one open to experiment.

Roughly speaking, around the turn of the century Peano, Jordan, Cantor, Borel, Stieltjes, and Lebesgue discovered and displayed with great subtlety the new properties of continuum, continuity and measurability. They have given a series of definitions of integration of increasing generality, and invented constructions and existence proofs for many strange mathematical objects which did not belong to the world of classical geometry and analysis but had to be accepted as a consequence of classical ways of mathematical reasoning stretched, as it seemed, to their limit.

The growing reaction against many counterintuitive discoveries led mathematicians to self–analysis centered around several basic problems: What is a mathematical proof? What meaning can be given to a statement about existence of a mathematical object? What is the status of mathematical infinity?

The outcome of this is well known. Fifty years of introspection were quite fruitful from the mathematical viewpoint: they produced mature mathematical logic, including theory of proof, theory of computability, and a clear picture of the hierarchy of expanding languages and axiom systems that mathematicians have had to adopt consecutively in their quest for mathematical truth.

In the meantime, physicists were engaged in a totally different quest. Planck's discovery of a quantum of action announced on December 14, 1900, initiated the quantum age. Physics needed sophisticated mathematics to formulate newly discovered non–classical laws, but new mathematics was of no help. Whatever was needed was hastily invented or reinvented: matrix algebra, spinors, Fock space, the delta function, the representation theory of Lorentz group. None of the pioneers (Bohr, Einstein, Pauli, Schrödinger, Dirac) needed the Lebesgue integral, or was interested in the cardinality of continuum. Logic interested them even less.

This does not mean that physicists had no philosophical preoccupations; in fact they had. But if mathematicians discussed the relationships between language and thought, physicists were troubled by the relation of language to reality. The basic problem confronted by the critics of classical mathematics was the inexpressibility of infinity, related to the inherently finitary syntactic structure of language. The basic problem confronted in the Bohr–Einstein controversy was the inexpressibility of quantum indeterminacy, related to the inherently classical semantics of language. Philosophy of mathematics and philosophy of physics almost completely lost contact with each other. Such ardent critics of the alleged inadequacies of contemporary research as Brouwer in mathematics and Pauli in physics shared not a single common idea. Mathematical criticism tended to become deeply autistic, while physical criticism strived to find better ways to express complex reality.[3]

A gap formed in traditional professional interactions as well. From the first successes of the quantum electrodynamics in the thirties until the renewed interaction in the sixties, mathematicians contributed almost nothing to the main physics research program of this century: Quantum Field Theory. Similarly, physicists payed no attention not only to mathematical logic (understandably) or analytical number theory (traditionally), but also to the emerging algebraic topology. Thirty years later, topology was to become the new common ground for the two communities. Somewhat paradoxically, mathematics gained from this renewed interaction more than physics: new invariants of three– and four–dimensional manifolds, quantum groups, quantum cohomology were its fruits.

[3] It is characteristic that G. H. Hardy's Rouse Ball Lecture [Hardy 1929] on "Mathematical Proof" delivered in 1928 does not even mention existence of quantum physics.

The following well known empirical observation fits well into the picture we have sketched. Whenever a fresh mathematical tool for understanding physics is needed, physicists are very quick at inventing new or transforming already existing *algebraic* formalism to deal with it. We have already mentioned Heisenberg algebra, spinors and Dirac delta function. One can add the Schwinger–Dyson equation (for an otherwise undefined path integral), the Berezin integral on supermanifolds and Witten's topological invariants expressed as path integrals of a topological QFT. All this constitutes only a small sample of inventions which are by now thoroughly absorbed and transformed into honest mathematics.

It is "only" when one has to deal with infinitary constructions, that is, limits of various kinds, that mathematicians do their job unassisted. According to Bourbaki's chapters on integration [Bourbaki 1974], mathematicians contributed to the theory of integral in the last century exclusively careful analysis of limits.

After the creation of the modern notion of a topological space and the discovery of limiting procedures basic to measure theory, the next major package of startlingly new infinitary constructions was introduced by Alexander Grothendieck with his treatment of homological algebra, derived categories and functors, topos and sites. But this is another story.

5. Discussion

Direct contact between mathematical and physical modes of thought more often than not creates a tension. The basic values are different, the accepted types of social behavior clash, time scales for a problem to keep attention of the public tend to be incommensurable.[4] In a remarkable piece of introspection, Dyson [1972] has shown how impenetrable the walls between mathematics and physics can be in one and the same mind. We would be much more tolerant to each other if we could discern in ourselves the two personalities so convincingly displayed by Dyson. A recent discussion (cf. [Jaffe and Quinn 1993, 1994]) shows the vulnerability of our community, when in a period of renewed fruitful interaction we try to harmonize our attitudes to what is and what is not a proof, what may and what may not be published, and who

[4]The relevant psychological difficulties are not often expressed in print. For an interesting recent reaction see S. Mac Lane's contribution in [Jaffe and Quinn 1994] of which we cite only one sentence: "Thus, when I attended a conference to understand the use of a small result of mine, I heard lectures about 'topological quantum field theory', without a slightest whiff of a definition; I was told that the notion had cropped up at some prior conference, so that 'Everybody knew it.' "

should be credited for what.

All of this is fortunately restricted to our social life. It seems that deep insights survive however we mess them up, and it is precisely the complementarity of mathematical and physical thinking that makes their interaction creative.

The crucial distinction between the ways we present our ideas in the last half of this century lies not so much in our attitudes towards a rigorous proof as towards exact definitions.

Mathematicians have developed a very precise common language for saying whatever they want to say. This precision is embodied first of all in *the definitions* of the objects they work with, stated usually in the framework of a more or less axiomatic set (or category) theory, and in the skillful use of *metalanguage* (which our natural languages provide) to qualify the statements. All the other vehicles of mathematical rigor are secondary, even that of rigorous proof. In fact, barring direct mistakes, the most crucial difficulty with checking a proof lies usually in the insufficiency of definitions (or lack thereof). In plain words, we are more deeply troubled when we wonder what the author wants to say than when we do not quite see whether what he or she is saying is correct. The flaws in the argument in a strictly defined environment are quite detectable. Good mathematics might well be written down at a stage when proofs are incomplete or missing, but informed guesses can already form a fascinating system: outstanding instances are A. Weil's conjectures and Langlands's program, but there are many examples on a lesser scale.

The etymology of the term *de–fin–itio* shows that its primary function is to set strict limits. In the course of a given study, we agree to consider only locally compact topological spaces satisfying the countability condition, only finite–dimensional Lie algebras, only coarse moduli spaces of stable algebraic curves and so on. If we fail to mention a relevant restriction in the course of presenting a professional seminar, we will be politely reminded about it. If we claim to having done anything serious, our work will be carefully scrutinized for all the necessary caveats.

Of course, our definitions are far from being arbitrary. One function of a good definition is to be a carrier of analogies between various situations, and to this end the cage of a definition must be of optimal size. For example, one can convincingly argue that by far the most important result of the group theory is exactly the definition of an abstract group and its action on a set, because it describes a structure reappearing again and again in geometry, number theory, probability, the theory of space–time, theory of elementary particles, and so on. The whole ideology of Bourbaki's treatise consists in representation of mathematics as a building supported by a strict system of good definitions (axioms of basic structures). And since a good definition is sometimes the work of generations of good mathematicians, the temptation to believe that we already know them all can be great.

To the contrary, an inexperienced reader of the most interesting physical papers is often left in a vacuum about the precise meaning of the most common terms. Physicists are undoubtedly constrained by their own rules, but these rules are not ours. What is a current algebra, a supersymmetry transformation, a topological field theory, a path integral, finally? They are very open concepts, and it is precisely their openness that makes them so interesting.

Here is what the history of our two metiers teaches: we cannot live without each other. At least for some of us, life becomes dull if it goes on for too long without contacts with good physics.

It is the interaction with a wildly differing set of values that counts most.

As a perceptive study by Hardy Grant [1995] shows, in terms of cultural history of Isaiah Berlin's variety, mathematics is a very classical endeavour. In fact, it is based upon a commonly accepted idea of truth and ways to achieve it, forming a stable system. The Romantic Revolution of a century and a half ago did not really influence mathematics mainly because there was little place in it for personal whims.

In this century romantics comes from physics: the vast expanses of the Universe, the wonderfully erratic behavior of the microworld, the observer's subjectivism and the power of the unobservable, the Big Bang, the Anthropic Principle, our in turn humble and megalomaniacal attempts to cope with irreverent Nature.

Mathematics supplies hygienic habits and headaches.

Bibliography

Atiyah (M.F.)

[1989] Topological quantum field theories, *Inst. Hautes Études Sci. Publ. Math.*, 68 (1989), pp. 175–186.

Blanchet (C.), Habegger (N.), Massbaum (G.) and Vogel (P.)

[1995] Topological quantum field theories derived from the Kaufman bracket, *Topology*, 34-4 (1995), pp. 883–927.

Bourbaki (N.)

[1974] *Élements d'histoire des mathématiques*, éd. rév. Paris : Hermann 1974. English transl. Heidelberg: Springer, 1994.

Candelas (P.), de la Ossa (X.), Green (P.) and Parkes (L.)

[1991] A pair of Calabi–Yau manifolds as an exactly soluble superconformal theory, *Nuclear Phys.*, 359 (1991), pp. 21–74.

Dyson (F.)

[1972] Missed opportunities, *Bull. Amer. Math. Soc.*, 78 (1972), pp. 635–652.

Feynman (R.P.)

[1948] The space-time approach to non-relativistic quantum mechanics, *Rev. Mod. Phys.*, 20 (1948), pp. 367–387.

[1988] *QED. The Strange Theory of Light and Matter*. Princeton Univ. Press, 1988.

Glimm (J.) and Jaffe (A.)

[1981] *Quantum physics. A functional integral point of view*. New York - Heidelberg - Berlin: Springer, 1981.

Grant (H.)

[1995] What is modern about "modern" mathematics?, *Math. Intelligencer*, 17-3 (1995), pp. 62–66.

Hardy (G.H.)

[1929] Mathematical proof, *Mind*, XXXVIII-149 (1929), pp. 1–25.

Jaffe (A.) and Quinn (F.)

[1993] Theoretical mathematics: toward a cultural synthesis of mathematics and theoretical physics, *Bull. Amer. Math. Soc.*, (N.S.) 29 (1993), pp. 1–13.

[1994] Response to comments on "theoretical mathematics", *Bull. Amer. Math. Soc.*, (N.S.) 30 (1994), pp. 208–211.

Reshetikhin (N.) and Turaev (V.)

[1991] Invariants of 3-manifolds via link polynomials and quantum groups, *Invent. Math.*, 103 (1991), pp. 547–597.

Witten (E.)

[1989] Quantum field theory and the Jones polynomial., *Comm. Math. Phys.*, 121 (1989), pp. 351–399.