IS OUR MATHEMATICS NATURAL?
THE CASE OF EQUILIBRIUM STATISTICAL MECHANICS
DAVID RUELLE

ABSTRACT. Is human 20th-century mathematics a natural or an arbitrary logical construct? Some insight into this ill-posed but fascinating question is gained from the study of mathematical physics. A number of ideas introduced by physical rather than mathematical necessity turn out later to be mathematically useful. (Equilibrium statistical mechanics has been a particularly rich source of such ideas.) In view of this the author argues that our mathematics may be much more arbitrary than we usually like to think.

1. Is our mathematics natural? The story goes that, when he reached heaven, Wolfgang Pauli requested to see his Creator, and asked Him to explain why the fine structure constant of electrodynamics has the value \( \alpha \approx 1/137 \). The Almighty went to the blackboard, and was busy writing formulae for a couple of hours. During that time Pauli sat, shaking his head and not saying a word. When finally the answer came: \( \alpha^{-1} = 137.0359 \ldots \), Pauli stood up, still shaking his head, took the chalk and pointed to an essential flaw in the calculation. I heard the story from Res Jost, and I wouldn't bet that it is completely authentic. Anyway, I think that many of us would like to ask some questions about physics and mathematics of Him who knows—when the opportunity arises. There are a number of obvious questions. For instance, about the consistency of mathematics: has He perhaps set up things, as Pierre Cartier suggests [1], so that the axioms of set theory are inconsistent, but a proof of contradiction would be so long that it could not be performed in our physical universe? Is this universe of ours the best of all possible worlds? Is it the only one of its kind, or could the fine structure constant be different from what it is? What kind of mathematics could be developed by intelligent beings living on a distant planet? Or in another universe with different physical laws?

Henri Poincaré once remarked that, for a question to make sense, one should be able to conceive of an answer which makes sense. This is not necessarily the case for the problems stated above. In fact, the problems which interest us most are often not easy to formulate satisfactorily. As a
consequence we usually settle for less, studying questions which make good sense mathematically, but may be philosophically somewhat futile.

In this Gibbs Lecture I would like to address a modest aspect of an ambitious question. Here is the ambitious question: *How natural is our human 20th-century mathematics?* I am willing to define mathematics as a logical construct based on the axioms of set theory. The question is: how much does the existing construct depend on human nature and condition? Those human contingencies and historical accidents should not change the truth value of theorems, but might dramatically influence the course taken by mathematical research, and the organization of the results obtained. I have stated the problem of the *naturality* of our mathematics in its ambitious generality, but I shall in fact restrict myself to a rather modest aspect of it. Before getting modest, however, let me make the bold suggestion that perhaps in a few decades we shall see what nonhuman mathematics looks like. I am not predicting the imminent arrival of little green men from outer space, but simply the invasion of mathematics by computers. Since the human brain is a sort of natural computer, I see no reason why the artificial variety could not perform better in the specialized task of doing mathematical research. My guess is that, within fifty or a hundred years (or it might be one hundred and fifty) computers will successfully compete with the human brain in doing mathematics, and that their mathematical style will be rather different from ours. Fairly long computational verifications (numerical or combinatorial) will not bother them at all, and this should lead not just to different sorts of proofs, but more importantly to different sorts of theorems being proved.

Be that as it may, I shall now return to human mathematics. Historically, the investigation of the physical universe in which we live has been all-important in the shaping of our mathematics. Geometry comes from the study of physical space, differential equations are linked with mechanics, and so on. But it is also clear that 20th-century mathematics now largely produces its own problems, and that physics is only a secondary source of inspiration. Bourbaki’s dream, and that of several generations of mathematicians, has been to find the natural *structures of analysis* and to develop them for their own merit. One can probably say that this dream has been the most powerful and fruitful source of inspiration for 20th-century mathematics. The role of physics, and now computer sciences, has been important but less central. One may thus feel that our mathematics has a rather natural *central core*, even though important branches may be influenced or motivated by applications to physics or computing. The many observed relations between different fields of mathematics support the idea of a natural central core. Little green men from outer space would have a similar central core for their mathematics; maybe it would be presented quite differently, but in their own language they would have a theorem saying that the image of a compact set by a continuous map is compact.

It is this confidence about the natural central core of mathematics which I would like to question in this lecture. My guess is that the central core of our mathematics and that of the little green men from outer space may not have much in common. We have confidence in the naturality of our mathematics
because of its unity: many relations are observed between different fields. But it is not the unity of mathematics which is doubted: the mathematics of little green men may have many relations with ours, even though it would not have much in common. To use a pedantic geometric image, think of different mathematical fields as marbles (of the same diameter to simplify things). We may have an arbitrarily large number of them with all pairs touching (provided we are in a space of sufficiently large dimension). It is thus not difficult to imagine our many marbles and their many green marbles all close together, yet completely disjoint.

Fine!—you may say—but how can one argue that our mathematics is not natural, without invoking computers of the future or little green men from outer space? What I would like to do is to appeal to mathematical physics. More precisely, I shall look for examples of mathematical ideas of physical origin which turn out to be mathematically natural and useful, although one would not have come to them easily if they had not been given to us, so to say, from outside. I claim that equilibrium statistical mechanics provides such examples.

But before I try to prove this fairly modest claim, I have to explain what mathematical ideas of physical origin are supposed to be. Understanding the universe in terms of mathematical constructs or laws is not a simple affair. Obviously, the laws of physics are not decreed arbitrarily by man; neither are they proclaimed unambiguously by Nature itself. In fact, it remains a mystery why Nature can be so well described by mathematical constructs (see Eugene Wigner's beautiful paper on *The unreasonable effectiveness of mathematics in the natural sciences* [2]). We shall not concern ourselves with this mystery, or discuss the epistemology of physics. All that will matter for us is that in some branches of physics—notably equilibrium statistical mechanics—some deep mathematical concepts arise which have been forced on us by the study of the laws of physics. It would have taken a very long time to reach these concepts in a mathematical study uninfluenced by physics.

2. Equilibrium statistical mechanics as a source of mathematical concepts. Mathematical physics consists of the analysis of particular mathematical idealizations for different classes of natural phenomena. In the worst cases this gives rise to a fragmented multitude of little mathematical problems of no particular interest. Often enough, though, the little problems beg for insertion into a wider theory, and this process of aggregation leads to a new mathematical field where the concepts are supplied by physics—often with compelling mathematical necessity.

A fact which is, at first sight, rather astonishing, is that the aggregation process just mentioned may unite unrelated areas of physics. Part of the explanation for this is sociological, and rests on the existence of a community of mathematical physicists. But another—essential—reason why the same new ideas serve to organize unrelated fields of physics is that these ideas are mathematically natural.

For instance, in recent years, the ideas and methods of equilibrium statistical mechanics have invaded relativistic quantum field theory to the extent of
practically causing a fusion of the two fields. [This is visible both at the level of the renormalization group ideas of Kenneth Wilson, and in the more rigorous work of constructive field theorists.] In a completely different direction, the concepts of equilibrium statistical mechanics have been found useful in the study of differentiable dynamical systems, and from there in the discussion of chaos and hydrodynamic turbulence. Note that the relation which this gives between statistical mechanics and fluid dynamics is purely mathematical: it has nothing to do with the fact that the fluid under study is also described—at the microscopic level—by (nonequilibrium) statistical mechanics.

Besides the contribution of equilibrium statistical mechanics to the study of differentiable dynamical systems mentioned above, there have been contributions to other mathematical disciplines. Remember that ergodic theory has its historical origin in statistical mechanics. Remember also that the statistical mechanical definition of entropy has been used by Claude Shannon to introduce the concept of information, which has in turn given the Kolmogorov-Sinai invariant of ergodic theory. In a completely different direction, the KMS equilibrium condition of quantum statistical mechanics has played an important role in the development of the Tomita-Takesaki theory of modular automorphisms of von Neumann algebras (see below).

At this point it is necessary to become a bit more explicit and specific. I shall start by a short general description of what equilibrium statistical mechanics is about,\(^1\) then show three examples of contributions of equilibrium statistical mechanics to mathematics.

Statistical mechanics deals with “large systems”, i.e., large collections of identical subsystems contained in a box \(\Lambda \subset \mathbb{R}^3\), in the limit where \(\Lambda\) becomes infinitely large. For commodity one commonly replaces \(\mathbb{R}^3\) by the “lattice” \(\mathbb{Z}^3\) (for instance by \(\mathbb{Z}\)). Typically, for classical systems, a compact set \(F\) and a finite positive measure \(m\) on \(F\) are given. A configuration of the system is described by a point \(\xi \in F\) for each \(i \in \Lambda\). A statistical state in \(\Lambda\) is simply a probability measure \(\rho_\Lambda\) on \(F^\Lambda\). To obtain \(\rho_\Lambda\) one chooses an energy function \(E_\Lambda: F^\Lambda \to \mathbb{R}\), and a number \(\beta > 0\) (inverse temperature); then one writes

\[
\rho_\Lambda(d\xi) = \frac{1}{Z_\Lambda} \exp(-\beta E_\Lambda(\xi)) \prod_{i \in \Lambda} m(d\xi_i),
\]

where \(Z_\Lambda\) is the partition function

\[
Z_\Lambda = \int \exp(-\beta E_\Lambda(\xi)) \prod_{i \in \Lambda} m(d\xi_i).
\]

Suppose that \(E_\Lambda\) is given for all finite \(\Lambda \subset \mathbb{Z}^3\), and satisfies an approximate additivity property: \(E_{\Lambda_1 \cup \Lambda_2} \approx E_{\Lambda_1} + E_{\Lambda_2}\) for large disjoint \(\Lambda_1, \Lambda_2\). Then it makes sense to study the limits of \(\rho_\Lambda\) when \(\Lambda \to \mathbb{Z}^3\); these are measures on \(F^{\mathbb{Z}^3}\), invariant under translations of \(\mathbb{Z}^3\), and called equilibrium states. They are the central objects of (classical) equilibrium statistical mechanics.

The fact that equilibrium states are natural objects has to do with equivalence of ensembles. This notion would take some time to explain (see [8, 9])

---

\(^1\)Mathematically oriented monographs on different aspects of equilibrium statistical mechanics include [3, 4, 5, 6, 7].
and that is probably why equilibrium statistical mechanics has been developed from physical considerations rather than from mathematical necessity.

The amount of randomness per unit volume in an equilibrium state \( \rho \) is a physically natural quantity corresponding to the thermodynamic entropy (Boltzmann). As indicated earlier, the physical entropy concept leads to the definition of the Kolmogorov-Sinai invariant of ergodic theory. The study of the dependence of equilibrium states on \( \beta \) is mathematically very difficult, but physically important because discontinuities of this dependence correspond to phase transitions. Different equilibrium states (corresponding to different \( \beta \)'s or to different interactions) are in general mutually singular measures. Physical necessity then forces us to study families of mutually singular measures, and the same physical necessity also gives hints on how to do it. The powerful methods thus developed can then be used for other problems, for instance the study of hyperbolic differentiable dynamical systems (using a correspondence with statistical mechanics established by “Markov partitions”).

What we have sketched is the statistical mechanics of classical systems. For quantum systems, the energy function \( E_\Lambda \) is replaced by an operator, integrations by traces, and an equilibrium state \( \rho \) is a state on a \( C^* \)-algebra rather than a probability measure.

3. Three examples and some reflections. I shall proceed with a brief description of three examples rooted in statistical mechanics and blooming—more or less—in mathematical theory.

The first example, the Lee-Yang circle theorem, is a beautiful failure.

**Theorem (Lee and Yang [10]).** Let \( a(\{i,j\}) \in \mathbb{R}, \) with \(-1 < a(\{i,j\}) < 1\), for \( \{ij\} \subset \{1,\ldots,n\} \). Write

\[
P(z) = \sum_{X \subset \{1,\ldots,n\}} z^{|X|} \prod_{i \in X} \prod_{j \notin X} a(\{i,j\}),
\]

where \( |X| = \text{card} X \) and the terms of order 0 and \( n \) are 1 and \( z^n \). Then all the zeros of \( P \) are on the circle \( |z| = 1 \).

In the application to statistical mechanics, \( P \) is the partition function \( Z_\Lambda \) for a ferromagnetic spin system and the position of the zeros of \( Z_\Lambda \) controls phase transitions. This theorem, conjectured on a physical basis, originally took some effort to prove. A later idea, due to Taro Asano, now permits a short proof (see Appendix). I have called this beautiful result a failure because, while it has important applications in physics, it remains at this time isolated in mathematics. One might think of a connection with zeta functions (and the Weil conjectures); the idea of such a connection is not absurd, as our second example will show. But the miracle has not happened: one still does not know what to do with the circle theorem.

The second example is the application of ideas of statistical mechanics to differentiable dynamics, made possible by the introduction of Markov partitions (Sinai [11, 12, 13], Bowen [14, 15]). Markov partitions transform the problems of ergodic theory for hyperbolic diffeomorphisms or flows into problems of statistical mechanics on the “lattice” \( \mathbb{Z} \). Among the many applications
of the method, let us mention Sinai’s beautiful proof [13] that hyperbolic diffeomorphisms do not necessarily have a smooth invariant measure.\(^2\) Also, since the geodesic flow on manifolds of negative curvature is hyperbolic, one has the possibility of studying zeta functions of Selberg’s type. Using Markov partitions, these zeta functions are expressed as certain sorts of partition functions, which can be studied by statistical mechanics. Thus one obtains for instance a “prime number theorem” for the lengths of closed geodesics on a compact manifold of negative curvature (not necessarily constant)—see [16].

Our third example pertains to the problem of extending to noncommutative algebras the concepts which are natural for commutative algebras. (Commutative algebras are intimately related to geometry, which is a rich source of natural concepts.) Just as classical statistical mechanics provides powerful techniques for handling measures (a problem of commutative algebras), quantum statistical mechanics is a rich source of inspiration for noncommutative algebras. Let me now describe the third example.

For a quantum system in a box \(A\), with energy operator \(E_A\), the time evolution of an operator \(A\) is given by

\[
A \mapsto \tau^t_A A = e^{iE_A t} A e^{-iE_A t}.
\]

On the other hand (taking the inverse temperature \(\beta\) equal to 1) we have a state \(\rho_A\) defined by

\[
\rho_A(A) = \frac{1}{Z_A} \text{tr} A e^{-E_A}, \quad Z_A = \text{tr} e^{-E_A}.
\]

Kubo, Martin and Schwinger remarked that there is a bounded continuous function \(F\) on \(\{z: 0 < \text{Im}\, z < 1\}\), analytic for \(0 < \text{Im}\, z < 1\) and such that for all real \(t\)

\[
\rho_A(B \cdot \tau^t_A A) = F(t), \quad \rho_A(\tau^t_A A \cdot B) = F(t + i).
\]

This situation persists when \(A \to \infty\): The equilibrium state \(\rho\) satisfies the KMS (Kubo-Martin-Schwinger) boundary condition with respect to the time automorphism \(\tau^t\). On the mathematical side, Tomita and Takesaki discovered and proved that if \(\rho\) is a faithful normal state on a von Neumann algebra \(M\), there is a unique one-parameter automorphism group \((\tau^t)\) of \(M\) such that \(\rho\) satisfies the KMS condition with respect to \((\tau^t)\). This unexpected connection between statistical mechanics and the theory of von Neumann algebras was found after about ten years of independent developments on the physical and the mathematical sides. (The connection was of course very beneficial for both sides). The story is excellently told by Huzihiro Araki [17], who stresses the key role played by a paper of Haag, Hugenholtz and Winnink [18]. The ensuing developments have led to a rebirth of the theory of von Neumann algebras, rebirth in which the work of Alain Connes has played a central role.

\(^2\)More precisely: There is a dense open set of \(C^\infty\) Anosov diffeomorphisms which do not have an invariant measure absolutely continuous with respect to the Riemann volume.
Our three examples were chosen somewhat arbitrarily among a number of others\(^3\) to show how different ideas originating in equilibrium statistical mechanics have fared differently in pure mathematics. Thus, in the third example the contribution of statistical mechanics was less essential than in the second, because of the independent development of the Tomita-Takesaki theory. In all cases, however, the impact of physical ideas was considerable. What does this tell us?

There is no question that, among the physical theories of this century, equilibrium statistical mechanics has been extraordinarily productive of deep mathematical ideas. We cannot say that we fully understand why. By comparison, nonequilibrium statistical mechanics has been mathematically rather unproductive, although the conceptual problems it raises are not less interesting to the physicist. As far as we know, it is entirely possible that nonequilibrium statistical mechanics will some day develop into a rich source of mathematical inspiration.

But let me come back to the original question of the "naturality" of mathematics. There are some natural ways in which mathematics develops: trying to solve questions which appear interesting, while using methods which appear feasible. Often enough, a long development leads to a result which now appears central, and a short cut is taken to this result, involving a change in point of view. This means that what is considered natural changes with time. For example, parallelism between mathematical theories (like "duality" in projective geometry) has been noticed long ago and used heuristically. But today one would find it ridiculous to derive two long series of parallel theorems, marvelling forever at their correspondence: one would instead try to define an isomorphism between the two theories and get rid (in effect) of one of them.

Let us now drag physics into the picture. Because of peculiar relevance criteria, what is natural from the point of view of physics is often quite different from what appears mathematically natural. The intrusion of physics therefore changes the historical development of mathematics. I have indicated that this intrusion of physics has been relatively modest in our century; it does however put into question the naturality of our mathematics. Different historical developments are conceivable, as indicated by the examples discussed above, and by many more which could be considered.

It may be time to attempt some cautious conclusions, remembering that the question of naturality of our mathematics is not a very well-posed problem. The influence of historical accidents should not be overstated: some concepts like those of natural numbers, or groups, would have to appear sooner or later in the development of human mathematics. Yet it is striking that some very good mathematical ideas have not been supplied by the internal logics of mathematical development, but have come from outside. Other external circumstances are conceivable, which would have led to different mathematics.

---

\(^3\) The recent definition of a "noncommutative Kolmogorov-Sinai invariant" by Connes, Narnhofer and Thirring \([19, 20]\) comes to mind, or the applications to relativistic quantum mechanics, for which see \([21]\).
How different? My own judgment is that human mathematics could be quite different from what it is.

I further think that whatever naturality our mathematics does have is not due so much to logical necessity as to the peculiar nature of the human mind. I mean the way in which our "logical thinking" is linked with visual intuition and tied to illogical "natural languages". I mean also our idiosyncratic liking for short formulations which we call "elegant", and for somewhat repetitive ways of arguing which we describe as "natural".

This is not to say that mathematics is an arbitrary construct. Of course not: it is a structured subject and in some sense it is nothing but structure. That structure was not made for us to understand, yet the human mind can grasp it: that is what makes mathematics so fascinating. We like to think of the discovery of mathematical structure as walking upon a path laid out by the Gods. But, as Antonio Machado wrote, maybe there is no path:

Caminante, son tus huellas
el camino y nada más;
caminante, no hay camino,
se hace camino al andar.

[Walker, your footsteps are the path, the path is nothing more; / walker, there is no path. / You make the path by walking.]

Appendix: Proof of the Lee-Yang Circle Theorem. Consider the class \( \mathcal{A} \) of polynomials of the form \( P_I(z_I) \) where

(i) \( I \) is a finite set.
(ii) \( z_I = (z_i)_{i \in I} \).
(iii) \( P_I(z_I) \) is separately affine in each of the card \( I \) complex variables \( z_i \).
(iv) If \( |z_i| < 1 \) for all \( i \in I \) then \( P_I(z_I) \neq 0 \).

Notice the following facts:

I. For real \( a \in [-1,1] \) the polynomial

\[
P_{\{1,2\}}(z_1, z_2) = z_1 z_2 + a z_1 + a z_2 + 1
\]

is in the class \( \mathcal{A} \). [Easy to see, by considering the involutive projective transformation \( z_1 \mapsto z_2 \) defined by \( P_{\{1,2\}}(z_1, z_2) = 0 \): the fixed points are on the unit circle.]

II. If \( I \) and \( J \) are disjoint and \( P_I(z_I), P_J(z_J) \) are in the class \( \mathcal{A} \), then the product \( P_{I \cup J}(z_{I \cup J}) = P_I(z_I)P_J(z_J) \) is in the class \( \mathcal{A} \).

III. Let \( I \cap \{\alpha, \beta\} = \emptyset \) and write

\[
P_{I \cup \{\alpha, \beta\}}(z_{I \cup \{\alpha, \beta\}}) = A z_\alpha z_\beta + B z_\alpha + C z_\beta + D,
\]

where \( A, B, C, D \) are polynomials in \( z_I \). Let \( \gamma \notin I \) and define the Asano contracted polynomial

\[
P_{I \cup \{\gamma\}}^*(z_{I \cup \{\gamma\}}) = A z_\gamma + D
\]

then if \( P_{I \cup \{\alpha, \beta\}} \) is in the class \( \mathcal{A} \), so is \( P_{I \cup \{\gamma\}}^* \). [Take \( |z_i| < 1 \) for \( i \in I \), then \( a z^2 + (B + C)z + D \) does not vanish for \( |z| < 1 \). The product \( D/A \) of its roots therefore satisfies \( |D/A| \geq 1 \), hence \( A z_\gamma + D \) does not vanish for \( |z_\gamma| < 1 \).]
THEOREM. Let $I$ be a finite set and choose $a({i,j}) \in \mathbb{R}$, with $-1 \leq a({i,j}) \leq 1$ for all $\{i,j\} \subset I$. Write

$$P_I(z) = \sum_{X \subset I} z^X \prod_{i \in X, j \notin X} a({i,j}),$$

where $z^X = \prod_{i \in X} z_i$; then $P_I$ is in the class $A$.

To see this, consider the polynomials $z_i z_j + a({i,j})(z_i + z_j) + 1$ (which are in $A$ by $I$). By successive multiplication of those polynomials, making an Asano contraction whenever the same variable appears twice, one obtains $P_I$. From II and III it follows that $P_I$ is in the class $A$.

The Lee-Yang circle theorem is an immediate corollary of the above results. Note that Asano treated a more difficult problem: he proved a quantum version of the Lee-Yang circle theorem. (See Asano [22, 23], and also Suzuki and Fisher [24], Ruelle [25].)

REFERENCES


**Institut des Hautes Études Sciences, 91440 Bures-sur-Yvette, France**