# The Effectiveness of Mathematics in Fundamental Physics

A. Zee
*Institute for Theoretical Physics*
*University of California*
*Santa Barbara*
*CA 93106, USA*

When I was a kid, so to speak, I came across and read Eugene Wigner's article[1] on "The Unreasonable Effectiveness of Mathematics in the Natural Sciences". I was impressed. Thus, I was rather pleased when Ron Mickens asked me, this many years later, to write a similar article.

Upon re-reading Wigner's article recently, I realized that Wigner had precious little to say, and perhaps necessarily so. I can hardly imagine that anyone could have much substantive to say about the question implied by the title of Wigner's article. Some of my colleagues may feel that, as a working theoretical physicist, I should not even be thinking about this question. But surely part of the fun of being a theoretical physicist is to be able to muse about such questions.

Furthermore, thirty or so years have passed since Wigner's article was published, and what we think of as physics and mathematics have shifted, at least in focus. It may be worthwhile to look at the question again. I want to say right from the start that I have neither a coherent theme nor a startling insight. Rather, I can offer only a disconnected series of observations, anecdotes, and musings. Anyway, what follows was known in my college days as "shooting the breeze".

Wigner's contribution lies, of course, in raising the question in the first place. In connection with Wigner's ability to ask seemingly profound questions, I may perhaps tell a little story. At around the time when I first read Wigner's article, I was an undergraduate. One winter day, I had to go back to the physics department after dinner to do some work. While I was eating dinner, it started to snow heavily outside. Trudging over to the physics building, I slipped and fell a couple of times. By the time I got there, I was literally covered with snow. As I staggered into the building, Eugene Wigner, with his heavy overcoat and hat on, was just about to go out. He looked at me carefully, and then he asked me, in that strangely solemn way that he had, "Excoose me, pleeze, izit snowing outside?"

Now surely that was a truly profound question if I ever heard one. The unreasonable effectiveness of visual inspection in comprehending reality? But while I did battle with such deep thoughts, Wigner had already walked out and disappeared into the howling storm.

A physics colleague once remarked to me that questions such as whether the effectiveness of mathematics in the natural sciences is reasonable or not have the curious property of being either incredibly profound or incredibly trivial. I am inclined to think that they are profound. Leaving that aside, let us try to understand and define the words in Wigner's question.

Effectiveness? What do we compare mathematics with in the effectiveness sweepstake? The trouble is that any alternative we can think of is so utterly ineffective by comparison. But how do we know that there isn't something more effective than mathematics? A prehistoric Wigner might have mused about the unreasonable effectiveness of magical incantations in understanding reality. Was there anything more effective than mathematics that we have not conceived of as yet by definition? Was Wigner asking about the unreasonable effectiveness of whatever we had found to be the most effective?

Indeed, I suspect that what Wigner means by mathematics may just be the entire world of quantitative notions. If so, then the question he raised may be profoundly trivial (or trivially profound). Of course, the bag of quantitative notions is more effective than the bag of qualitative notions, if only because quantitative notions are more precise and compact.

How do we measure reasonableness? According to some sort of community standards? Evidently, since physicists assume that space and time are continuous, and since the assumption is based on well-verified observations, it stands to reason that the mathematics of continuously differentiable functions would be effective. In this view, since the basic mathematical concepts, such as those of geometry, are abstracted out of our experience of the physical world, mathematics ought to be effective. Proponents of this view point to the apparent fact that those areas of modern mathematics least rooted in "everyday" experience tend to be irrelevant to physics.

Perhaps the most dramatic counterexample to this view is the emergence of the complex number in quantum physics. Why, indeed, should the microscopic world be described by complex numbers? It is really rather mysterious.

Coming back to the effectiveness of differentiable functions in describing the physics of continuous spacetime, I should think that if at some distance scale, we were to discover that spacetime is actually discrete, then mathematics (as understood by Wigner, say) would not be effective at all. Indeed, mathematics is not particularly effective in areas such as lattice gauge theory. Not much can be done besides letting the computers go for it.

I want to move on and spend the most time defining "mathematics". Wigner had devoted two sections to "What is mathematics?" and "What is physics?". Rather than repeat what he said, I want to distinguish between mathematics and, for lack of a better term, arithmetic. This distinction is frankly and intentionally touched with a measure of snobbery. ("What you call mathematics is merely arithmetic to me.") Twenty years after Wigner's article appeared, R. W. Hamming wrote an article[2] titled "The Unreasonable Effectiveness of Mathematics". Period. I didn't care much for Hamming's article, because it seemed to me that Hamming, drawing upon his engineering experience, was mostly talking about what I call arithmetic.

Alright, what is the distinction? I would say that mathematics is whatever a reasonably brilliant physicist, defined for the purpose here as someone significantly smarter than I am, could not work out in a finite amount of time, by following more or less straightforward logic. (I will leave it to you to haggle over how long a finite amount of time is.) Everything else is arithmetic. For instance, I probably could have figured out the properties of the solutions of Legendre's equation. All that stuff about Legendre polynomials is definitely arithmetic. On the other hand, the fact that there are only three cases in which higher dimensional spheres can be mapped non-trivially onto lower dimensional spheres, namely $S^3 \to S^2, S^7 \to S^4$, and $S^{15} \to S^8$, that I call mathematics.

Whether I call something arithmetic or mathematics depends to some extent on how we look at it. If we recognize Legendre polynomials as having to do with representations of the rotation group, that indicates some understanding of the structural properties of rotations. In short, I associate mathematics with structural or global understanding, and arithmetic with computation.

Echoing a fairly widespread arrogance of the physicist, Feynman once said that had God not created mathematicians, physics would have been delayed by about a week. (Had complex numbers not been invented by the 1920s, would the development of quantum mechanics have been delayed by significantly longer than a week?) According to Feynman, physicists would have invented what they needed, and the rest, as far as he was concerned, should not have been bothered with in the first place. Feynman's attitude, of course, represents a long tradition in physics. Until the mid-1970s, I would have been inclined to agree with Feynman, but with the advent of superstring theory, and for about a decade before that, truly profound mathematics had started coming into physics, with an intensity that was last seen with the arrival of group theory into quantum physics. But before these developments, it seemed as if physicists could really follow Feynman and have nothing to do with the "pure" mathematicians.

In a way, Feynman was right. We reached the grand unified theory in the 1970s using a minimal amount of mathematics. Some fairly elementary group theory, that's about it. For Pete's sake, the grand unified theory! The theory that unifies three of the fundamental interactions! We have unravelled a big piece of Nature's innermost secrets about how the world is put together without using anything a mathematician would call mathematics. Indeed, the creators of the grand unified theory, and most of the particle physicists of the 1970s, were very Feynmanesque in their disdain for mathematics. Incidentally, Feynman once told me, while we were watching some show, that fancy-shmancy mathematical physics as applied to physics is not worth "a bottle of piss-water".

I like to think of the development of particle physics since the mid-1970s (say) as breaking the shackles of Feynman diagrams. I believe that Feynman diagrams, with all the brilliant simplicity that they incorporate, had too long an influence, and ultimately an unhealthful influence, on particle physics. In a quantum field theory course I took in graduate school, the professor told us that field theory is defined as the totality of Feynman diagrams. The set of diagrams defines a unitary, analytic, and Lorentz invariant theory. All those manipulations in the standard canonical development, such as commuting field operators, are so fraught with delicacies that they are to be regarded as props needed to derive the Feynman rules. Once the rules are determined, quantum fields are to be thrown away.

In this climate, there was indeed no need to learn any mathematics. This view was swept away by the advent of concepts such as solitons and instantons. Particle physicists had to learn about such fancy-shmancy mathematics such as topology.

The ten-year gap between the understanding of spontaneous symmetry breaking and of solitons can, in my opinion, be attributed to the constraining influence of Feynman diagrams. Even in the 1970s there were many people who prefer to describe spontaneous symmetry breaking as the disappearance of diagrammatic lines into the vacuum. The notion of fields as real, as something that we can knead into twisted lumps, was quite revolutionary.

Ironically, the formalism that came into fore, namely the path integral formalism, was also developed by Feynman. That represents a real tribute to Feynman.

With the shackles of Feynman diagrams broken, Feynman's view on mathematics also started to fade. A younger generation of particle physicists felt increasingly at ease with modern mathematics. There was a fundamental shift in outlook towards mathematics, and with the advent of superstring theory around 1983 or so, the trend has accelerated. Today, much of the research in superstring theory is really research into mathematical structures, of a degree undreamed of by Wigner.

With this brief review of how the attitudes of theoretical physicists have changed over the last thirty years or so, let me now come to some observations about the role of mathematics in fundamental physics. You may have already noticed my restriction to "fundamental physics" in my title. I invented this term some years ago to replace the outmoded "particle physics" (or even worse "high energy physics"). Perhaps somewhat tautologically, I define a fundamental physicist as someone who is interested in discovering something fundamental about the physical world. Fundamental physics and particle physics overlap to a large extent, but neither contains the other. The definition is broad enough to include some condensed matter physicists who are interested in understanding the "global" properties of strongly quantum many-body systems. Were I to maintain that arrogant tone I used in distinguishing mathematics from arithmetic, I could have defined a fundamental physicist as someone who tends to use mathematics rather than arithmetic. Anyhow, let's go on.

I believe that the following is a true and somewhat mysterious fact: deeper physics is described by deeper mathematics. Consider the Schrödinger operator versus the Dirac operator. The mathematical structure underlying the Dirac operator is much richer and deeper than the structure underlying the Schrödinger operator. We would expect so since the Schrödinger equation is an approximation of the Dirac equation. Associated with the Dirac operator is real mathematics.

This point was underlined for me quite strikingly some time ago when a colleague and I were studying a condensed matter physics problem of a non-relativistic electron hopping on a two-dimensional lattice in the presence of quantized magnetic flux. The problem has nothing to do with relativistic physics and the Dirac equation. To determine the energy $E$ of the electron as a function of its momentum $p$, we have a completely standard and straightforward problem of finding the eigenvalues of some $n$ by $n$ matrix. For all but the simplest cases, one would have to go to the computer and crunch some numbers. Nothing inspiring or mathematical about it. However, suppose we are not interested in the precise function $E(p)$ but in determining the number of zeroes, that is, the number of places in $p$-space where $E$ vanishes. Around such a point $p^*$, we can expand and in general $E$ will depend linearly on $(p - p^*)$: schematically $E \simeq a(p - p^*)$. With a suitable shift and scale change in the definition of $p$, we see that the behavior of the electron in that region of momentum space is described effectively by a Dirac equation. What is remarkable is that the entire mathematical edifice of index theorems and winding numbers can now be brought to bear on finding out how many such $p^*$s there are.

The point is that to determine the function $E(p)$ or even to determine the locations of the $p^*$s is a job in arithmetic. These quantities depend on the details of the Hamiltonian. Change the Hamiltonian slightly and we expect the value of a given $p^*$ to shift. In contrast, the mathematics tells us that the number of $p^*$s is invariant as long as the overall structure of the Hamiltonian remains unchanged. In other words, we have the concept of a topological invariant here. Mathematics is effective in giving us global and structural understanding but not in solving computational problems.

I once remarked that more mathematics is associated with the Dirac operator than with the Schrödinger operator to a conference of philosophers interested in physics. Somebody in the audience objected vociferously, "Just count the number of mathematical papers written on the Schrödinger operator!" he said. The confusion here is between usefulness and beauty, so to speak. The Schrödinger equation is useful to a much larger group of physicists than the Dirac equation. Of course people would have devoted a great deal of energy to unravelling the properties of the Schrödinger equation.

Even if we were to focus on the Schrödinger equation, not all Schrödinger problems are created equal. Consider the Schrödinger problem associated with the Stark effect and the Schrödinger problem of a particle moving on a sphere around a magnetic monopole. The latter is associated with deep mathematics, the former is not. But, some of you may think, the former is at least useful, the latter is not, since the monopole may not even exist. In fact, precisely because of the deep mathematics associated with the monopole problem, it has been of central importance in recent developments in theoretical physics and has popped up all over the map. I will mention only a few examples: Berry's phase, Polyakov's instanton in $(2+1)$-dimensional compact gauge theory (a theory which may be relevant to high temperature superconductivity), Haldane's treatment of the antiferromagnetic spin chain, and the movement of holes in a ferromagnetic background. None of these problems has anything to do with the magnetic monopole *per se* but they all share the same underlying mathematical structure.

My story about finding the zeroes also illustrates a point known to every practicing physicist, namely the importance of asking the right question. There are several connotations to the word "right". Obviously, we want to ask physically relevant questions. But we also want to ask questions for which mathematics is effective.

Perhaps sadly, the importance of asking the right question has diminished in some areas of physics due to the availability of computers. Thus, in the problem mentioned above, one can simply compute everything numerically. To some extent, arithmetic can replace mathematics. But inevitably, arithmetic cannot provide the understanding brought by mathematics.

In the thirty years since Wigner's article, we have seen the computer become a major force in theoretical physics. The computer has extended, in a sense, the very domain of theoretical physics. Such fields as chaos and nonperturbative quantum chromodynamics would have been essentially impossible without the computer. At the same time, the computer has pronounced on the subject of Wigner's article: mathematics is not particularly effective in physics, if we define physics as the collection of problems and situations considered by the community of physicists.

An important role played by mathematics is in limiting the possibilities physicists have to consider. An example is the exhaustive classification of Lie algebras. This is obviously of great importance in the development of grand unified theories, for instance. I think that Feynman is wrong here about how physicists can just invent the mathematics that they need. I feel that physicists can probably work out the theory of a specific group, SU(5), say, but the reasoning that allows one to say "Here are all the possible Lie algebras and groups, folks!" is peculiarly mathematical. Actually, the reasoning involved, once it has been invented of course, is not particularly difficult to follow, but it carries that peculiar quality known as mathematical insight.

Of course, there are physicists around, some of the young string theorists, for example, who probably could have worked out the complete classification of Lie algebras. But then these people could have easily become mathematicians as well. Feynman's crack only makes sense if there are distinct personality types, so that out of two persons, equally intelligent according to some measure, one can only be a physicist, the other a mathematician. From my own observations, I believe that that's true to some extent. Many great physicists would have been hopeless as mathematicians.

In truth, there have been many examples of higher mathematics discovered independently by physicists. There is a well-known story by Molière about a gifted but unschooled writer. When a friend complimented him on his prose, our writer was puzzled until he learned that what he was writing was called prose. A physicist colleague of mine was fond of asking mathematically more sophisticated friends in a mocking tone, "Tell me, have I been writing prose?"

It often happens that a discovery in physics is actually associated with a wealth of mathematical structure undreamed of at the time of the discovery. (This is the point I made earlier about the magnetic monopole problem.) Dirac certainly could not have imagined all the mathematics associated with the Dirac operator. A wonderful example in particle physics is the chiral anomaly. It was first discovered in the late 1960s as an oddity: an explicit Feynman diagram calculation had shown that an alleged theorem derived by naively manipulating quantum fields was incorrect. (In fact, already starting in the early 1950s various people had stumbled upon the chiral anormaly in one form or another without recognizing it.) Over the last twenty some years, the chiral anomaly has had a totally amazing habit of popping up in connection with all sorts of major theoretical developments. These developments include the renormalizability of gauge theories, path integral measures, instantons, fractionization of quantum numbers, induced proton decay, winding numbers and intersection numbers, selection of a suitable superstring theory, just to mention some examples. The reason for this remarkable ubiquity is that the chiral anomaly turned out to be associated with a deep mathematical structure rooted in topology and geometry.

It does not follow, however, that objects of great interest to physics is necessarily associated with deep mathematical structures. Consider quantum field theories. In the modern view, a quantum field theory is defined by some sort of functional integral of an integrand equal to the exponential of the action times the imaginary unit $i$. Apparently, it is just some functional integral out of many possible functional integrals. At least thus far, physicists have not discovered any particularly deep mathematics associated with this integral. In contrast, a part of the integrand, namely the action, often has interesting mathematical properties. (An example is the pure Yang-Mills action.) But in all important cases, knowing some properties of the integrand is not of much help in understanding the properties of the integral. This is well known in statistical mechanics, for example.

Mathematics sometimes popped up uninvited in a physical situation. An example is that of an electron moving in a plane under a uniform magnetic field. The lowest energy quantum wavefunctions turn out to have the form $f(z)e^{-|z|^2}$ where $f(z)$ is *any* holomorphic function and where $z \equiv x+iy$ is the complex coordinate on the plane. Laughlin was able to develop an essentially complete theory of the fractional Hall effect by constructing a variational many-body wavefunction out of these wavefunctions. The structural properties of the theory are made apparent by invoking analyticity theorems at every turn. One has to repeatedly argue along the line "Such and such must have this particular form because of analyticity". *A priori*, the fractional Hall effect poses a formidably difficult problem in many-body dynamics and a theory as complete as Laughlin's would appear to be out of the question.

Remarkably, the wavefunction has the above-mentioned holomorphic form only in a certain gauge (out of an infinity of possible gauge choices). Indeed, one might not have the insight to write the wavefunction in terms of $z$ at all (but instead, in terms of the usual $x$ and $y$). Conceivably, one can still develop the same theory. After all, at every step, the equations can be gauged, transformed and re-written in terms of $x$ and $y$. The structural properties of the theory would then be totally obscured.

The preceding illustrates the well-known fact that often, using the right representation can be most of the battle.

This brings us to another fact known to all practicing physicists: the effectiveness (should we say reasonable or unreasonable?) of notation in doing physics. At the simplest but yet a profound level, algebra was invented when someone introduced the "notation" of using letters to represent quantities. In doing physics, we all have our favorite notations to the point that we can barely tolerate an unfamiliar notation. The human mind is a creature of habit. We are used to $m$ for mass and $T$ for temperature, and that's that. Some years ago, a distinguished particle physicist used the letter $\pi$ as an index (for example, for the $\pi$th component of momentum). His papers, which are already quite difficult to read, appeared all that more difficult.

I was told that Maxwell used to write out the components of the electric and magnetic fields starting with $E$ for the first component of the electric field, thus $E, F, G, H, I$ and $J$. (This is why the magnetic field is called $H$!) Whether or not this story was invented I do not know, but just imagine doing a standard problem in $E\&M$ using this notation! The introduction of indices is a truly neat trick. It has also been said that the repeated index summation represents one of Einstein's greatest contributions to physics.

In these examples, a better notation represents heightened efficiency in the sense of the accountant. But often a better notation implies a deeper understanding of the subject. For instance, Dirac's bra and ket notation underlines the fact that we do not have to specify the representation of a state vector in Hilbert space.

Indeed, there are entire topics of mathematics that amount essentially to a better notation. Take differential forms, for example. When I was a freshman, John Wheeler decided, as an experiment, to teach the introductory physics course from "the top down". (Thus, relativity and quantum mechanics were discussed first so that classical physics can be obtained as a "trivial" approximation. Incidentally, the experiment was not repeated the next year.) We were taught electromagnetism using differential forms: "indices without indices" (part of this discussion, using "egg-crates", later appeared in a well-known text on gravity co-authored by Wheeler). Needless to say, we were totally mystified. What was worse was that I, and probably others as well, developed a total distaste for differential forms. It appeared to me useless, since in any specific problems, we eventually had to write out the components of the differential form anyway. For years, I resisted differential forms even as several well-meaning colleagues tried to "teach" me the notation. But about seven years ago, when I was working on anomalies in higher dimensions, I suddenly realized that I could not live without differential forms. If you doubt this, just try solving for $w$ in an equation like trace $F^n = dw$ by writing out everything with indices. You will literally drown in a sea of indices. (Here $F$ is the Yang-Mills gauge field 2-form.)

The real advantage of differential forms is not so much that it saves us from writing out an endless streams of indices, but that it makes clear the geometrical character of various physical quantities. For example, in the magnetic monopole problem, the gauge field 2-form allows us to think of the gauge field $F$ as a single geometrical entity. The writing of $F$ in terms of its components, in contrast, requires commitment to a definite coordinate choice and splits a simple geometrical concept (namely, the concept of area) into an unrecognizable mess. As another example, in that problem mentioned in the preceding paragraph, physicists had long worked out, by using arithmetic, what $w$ is for the case when $n = 2$. But the recognition that as a form, (trace $F^n$) is closed but not exact conveys a truly deeper understanding.

The trouble was that when Wheeler was trying to convince me of the beauty of differential forms I was trying to master such problems as calculating the electrostatic field of a charged disk, in other words, problems that have about as much underlying geometrical structure as stock pricing analysis. There is no way that differential form can manifest its power in problems that call for only arithmetic.

(This suggests another "definition" of arithmetic versus mathematics: the electromagnetic potential around an electric charge = arithmetic, while the electromagnetic potential around a magnetic charge = mathematics.)

In recent years, various topological concepts such as linking and intersection numbers have entered into physics. Again, they can be written compactly and naturally in terms of a differential form with its underlying geometrical properties.

This story illustrates that mathematics is often too powerful for the physics. Differential form is too much for doing electromagnetism. But when you are tickling non-Abelian gauge theories in higher dimensional spacetime, then differential forms become indispensable.

My earlier attitude towards differential form is typical of the practicing physicist: I'm not gonna learn this stuff unless I can use it for something. My attitude towards fiber bundles, for example, remains at that stage. I have yet to encounter a physics problem in which fiber bundles would help me sigificantly, but I have no doubt that I will eventually. The mathematical concept expressed in fiber bundles strikes me as universal and natural and at some point it is going to seduce me for sure.

Fiber bundle provides an example in which it is useful just to know the words. They serve as pegs on which we can hang our physical concepts, so to speak. Often, they work as mnemonics. For instance, in the fundamental problem of a charged particle moving on a unit sphere around a magnetic monopole, words like sections, while of no actual help to us in solving the problem, remind us that the wavefunction is to be solved on separate patches and then joined together by gauge transformations. In recent years, physicists have used homotopy groups by and large in the same way, as mnemonics more than anything else.

Let us go back to the comparison between the Schrödinger and the Dirac operator. What we gave up in going from the Dirac operator to the Schrödinger operator is of course symmetry: Lorentz symmetry is broken down to rotational symmetry. Lately, I have been particularly struck by the awkwardness of non-relativistic equations when compared to their relativistic counterparts. I was trained as a relativistic physicist, but in the last year and a half I have been working on condensed matter physics. At first, my collaborator had to point out to me constantly that I had erroneously written down a relativistic equation. With a sigh, I would trudge through the non-relativistic form. It would invariably turn out to be much more tedious to deal with.

Deeper mathematics is associated with more symmetrical structures. In 1960, when Wigner wrote his article, the laws of the microscopic world looked rather asymmetric. We now know that those laws are merely phenomenological approximations to deeper laws, which are in fact symmetric. Symmetry has turned out to be a central organizing principle in Nature's design. Indeed, the story of fundamental physics in the last quarter of a century or so has been the profound discovery that as we study Nature at ever deeper levels, Nature exhibits ever larger symmetries.

I have told this story in considerable detail elsewhere.[3] Here I will merely emphasize that it does not have to be such that Nature's laws become more and more symmetric at deeper and deeper levels. For instance, there was a perfectly viable theory of the weak interaction in which the phenomenological Fermi theory was due to the exchange of a pair of scalar particles. Nature could have been designed so that the weak interaction would not be connected to the electromagnetic interaction at all.

Indeed, I think that we can raise the question of "the unreasonable effectiveness of symmetry considerations in understanding Nature". Why should symmetry dominate Nature at the fundamental level? Does the very fact that Nature becomes ever more symmetrical imply that there is a design? Einstein once said that the most incomprehensible thing about the world is that it is comprehensible. *A priori*, we could have lived in a chaotic universe whose working is beyond our comprehension. I have speculated on the philosophical issues raised by these questions in a recent article,[4] and so I will concentrate on the relationship between symmetry and mathematics here.

Symmetry and mathematics are closely intertwinned. Structures heavy with symmetries would also naturally be rich in mathematics. And so if it is indeed true that Nature's design becomes more symmetrical as we probe deeper and deeper, mathematics should be ever more effective.

Let me come back to the distinction between arithmetic and mathematics. In a broad sense, this split is mirrored by the split between dynamics and kinematics in physics. The application of fancy mathematics to physics often amounts to the erection of a kinematical framework within which we can ask dynamical questions. Mathematics is then often not particularly effective at this stage, and arithmetic has to be called upon.

As a specific example, I can refer to the recent discussion of, for lack of a better term, what may be called Chern-Simons theory, with its multifarious possible implications for subjects ranging from topological field theory with its connection to string theory and quantum gravity to high temperature superconductivity. The discussion can be wonderfully mathematical with fancy terms like Hopf terms and braid groups bandied about, but when it comes to actually understanding high temperature superconductivity we have to confront a "real-life" physics problem of working out the statistical mechanics of a liquid of particles with fractional statistics. What is the free energy of this liquid? What are its elementary excitations? Does it behave as a superfluid? Fancy math ain't gonna tell us nothin. Only physical insight and arithmetic will.

In the physics community, people are often involved in value judgment, talking about whether the problem so and so has solved is easy or hard. But in deciding whether or not to be impressed by a colleague's work, people tend to be impressed by fancy mathematics. Paradoxically, problems for which fancy mathematics are effective are often kinematical and hence easy. To quote an example perhaps of little physical importance, I recall that in constructing a Chern-Simons theory of membranes, my collaborator and I were guided at every step by the underlying mathematics of Hopf map, and we knew that things must work out in a preordained way (for instance, that quarternions must enter). Anyone who has worked on a physics problem with a heavy mathematical rather than arithmetical component must have had the feeling that the mathematics has a life of its own and can literally pull one along.

I speak of this split between arithmetic and mathematics from experience as I have worked on both types of physics problem. Perhaps somewhat strangely, I am attracted to both arithmetic and mathematics.

Next, I would like to mention a pervasive feeling among theoretical physicists expressed by the noble sentiment that "If the physics I am working on reveals an unexpectedly rich mathematical structure, then the physics must be correct." We all know that there had been some spectacular confirmations of this hypothesis: Einstein's theory of gravity and

Dirac's theory of the electron, for instance.

This argument is now invoked by some string theorists, and it is certainly true that the mathematical structure hidden in some apparently unprepossessing action describing a string is nothing short of incredible. String theory may well be right, but should we buy this argument? There is a nagging feeling among some people it is no coincidence that the structures studied by fundamental physicists are also precisely those structures favored by mathematicians. The string worldsheet is two-dimensional on which complex numbers and hence analytic functions naturally live. The action may be viewed as a conformal field theory, and all sorts of nice mathematics follow. From the point of view of a naive physicist, it would appear natural to study blobs, rather than strings, if we are going to study extended structures at all. Alas, the blob worldthing is a nasty place where no self-respecting analytic function or conformal field would dare set foot. There does not seem to be a decent mathematical structure at all. Does this mean that string theory is right?

Of course, it may be a waste of time to muse about such things. If I have the strength, I ought to be working on string theory instead. But the preceding discussion has brought us to what I call the dartboard theory of theoretical physics. In the mid-seventies, when there was a proliferation of models of the electroweak interaction, a distinguished experimentalist remarked to a group of us theoretical physicists that theorists are just throwing darts randomly, one of the darts is bound to land, and the wrong theories are just forgotten.

All the theorists who heard this remark were of course outraged, and I think rightly so. The textbook description of the development of physics as a competition between theories does not apply, for the most part, to fundamental physics. At any time, there is usually not a choice between theory $A$ and theory $B$. Rather, the choice is between a prevailing theory and nothing. We do not have the luxury of choosing between string theory and some other theory. Neither was gauge theory competing with some other theory during the 1970s.

Are mathematicians throwing darts randomly at physics? Out of the wealth of structures studied by mathematicians, isn't it reasonable that some of them are bound to be effective in understanding the physical world?

The influx of mathematics into particle physics over the last few years can only be described as a tidal wave. If you have not followed the development of string theory, let me give you a calibration. In 1984, a theoretical physicist who had a comfortable familiarity with such concepts as coset spaces, homotopy groups, homology sequences, and exceptional algebras would have been regarded by his colleagues as mathematically sophisticated. Some four years later, that same person would be despised by string theorists as a hopelessly unschooled mathematical ignoramus.

Is so much mathematics good for physics? I have no idea. The proof is of course in the eating: we will have to see if string theory can explain the world. Meanwhile, the tidal influx of mathematics is perfectly reasonable. In exploring the physics of the Planck scale, physicists are so far removed from any experimental moorings that mathematics can be our only guide, in a way that Wigner could not possibly have imagined.

In connection with the role of mathematics in physics, I am fond of telling the story[3] of Faraday and Maxwell. Because of his up-from-rags background, Faraday had a self-admitted blind spot – mathematics – and he was unable to transcribe his intuitive notions into precise mathematical descriptions. Just the opposite, Maxwell, scion of a distinguished family, received the best education, in mathematics and in everything else, that his era could provide. But before he began his investigations, Maxwell resolved "to read no mathematics on the subject (of electricity) till I had read through Faraday's *Experimental Researches on Electricity*". Indeed, he considered Faraday's mathematical deficiency an advantage. He wrote: "Thus Faraday ... was debarred from following the course of thought which had led to the achievements of the French philosophers, and was obliged to explain the phenomena to himself by means of a symbolism which he could understand, instead of adopting what had hiherto been the only tongue of the learned."

By "symbolism", Maxwell was referring to Faraday's "lines of force". Earlier, Maxwell had said that "the treatises of (the French philosophers) Poisson and Ampère (on electricity) are of so technical a form, that to derive any assistance from them the student must have been thoroughly trained in mathematics, and it is very doubtful if such a training can be begun with advantage in mature years." Well, I am sure that physicists "of mature years" can all empathize with what Maxwell said.

The story of Faraday and Maxwell is interesting particularly because it is not clear what moral it offers. I think that we are agreed that intuition in the grand tradition of Faraday has been of utmost importance in the development of physics. On the other hand, when you are wandering around in Planckland, what intuition can you possibly have? Let us not forget that Maxwell could probably not have been able to derive the propagation of light without using the methods of the French philosophers, namely differential equations. (Mathematics to him, but arithmetic to us.)

I like to close this musing about the effectiveness of mathematics, reasonable or otherwise, by telling another anecdote.[5,6] A lady who knew Einstein in her youth told me that once, on a brilliant spring day she and Einstein walked into a garden blooming with flowers. They stood looking at the scene in silence. Finally, Einstein said, "We don't deserve all this beauty."

Physics is a beautiful subject made all the more beautiful by the effectiveness of mathematics. Is it reasonable to think that we deserve all this beauty?

## References

1. E. P. Wigner, *The Unreasonable Effectiveness of Mathematics in the Natural Sciences*, *Commun. Pure Appl. Math.* **13** (1960) 1–14.
2. R. W. Hamming, *The Unreasonable Effectiveness of Mathematics*, *Am. Math. Mon.* **87** (1980) 81–90.
3. A. Zee, *Fearful Symmetry* (Macmillan Publishing Company, 1986).
4. A. Zee, *Symmetry in the Ultimate Design*, to appear in a book edited by R. Kitchener.
5. S. Asker, private communication.
6. A non-physicist friend to whom I told this story immediately interpreted Einstein's reaction in terms of the collective guilt of physicists in connection with nuclear weapons. I prefer to interpret it at a deeper level, in connection with the discoveries of ever larger symmetries in Nature's design.