

MISSED OPPORTUNITIES¹

BY FREEMAN J. DYSON

It is important for him who wants to discover not to confine himself to one chapter of science, but to keep in touch with various others.
JACQUES HADAMARD

1. Introduction. The purpose of the Gibbs lectures is officially defined as “to enable the public and the academic community to become aware of the contribution that mathematics is making to present-day thinking and to modern civilization.” This puts me in a difficult position. I happen to be a physicist who started life as a mathematician. As a working physicist, I am acutely aware of the fact that the marriage between mathematics and physics, which was so enormously fruitful in past centuries, has recently ended in divorce. Discussing this divorce, the physicist Res Jost remarked the other day, “As usual in such affairs, one of the two parties has clearly got the worst of it.” During the last twenty years we have seen mathematics rushing ahead in a golden age of luxuriant growth, while theoretical physics left on its own has become a little shabby and peevish. So I am forced to give this lecture an emphasis different from that intended by the founders. Instead of talking about “the contribution that mathematics is making to present-day thinking” in my field, I shall talk about the contribution that mathematics ought to have made but did not. I shall examine in detail some examples of missed opportunities, occasions on which mathematicians and physicists lost chances of making discoveries by neglecting to talk to each other. My purpose in calling attention to such incidents is not to blame the mathematicians or to excuse the physicists for our failure in the last twenty years to equal the great achievements of the past. My purpose is not to lament the past but to mould the future.

It is obviously absurd for me to imagine that I can mould the future with a one-hour lecture. The fact that Hilbert in 1900 [1] and Minkowski in 1908 [2] succeeded in doing it does not give me any confidence that I can do it too. But at least I have learned from Hilbert and Minkowski that one does not influence people by talking in generalities. Hilbert and Minkowski gave specific suggestions of things that mathematicians and physicists could profitably think about. I shall try to follow their

¹ Josiah Willard Gibbs Lecture, given under the auspices of the American Mathematical Society, January 17, 1972; received by the editors January 17, 1972.

style. I shall try to convince you by examining actual cases that the progress of both mathematics and physics has in the past been seriously retarded by our unwillingness to listen to one another. And I will end with an attempt to identify some areas in which opportunities for future discoveries are now being missed.

2. Digression into number theory. I begin with a trivial episode from my own experience, which illustrates vividly how the habit of specialization can cause us to miss opportunities. This episode is related to some recent and beautiful work by Ian MacDonald [3] on the properties of affine root systems of the classical Lie algebras.

I started life as a number theorist and during my undergraduate days at Cambridge I sat at the feet of the already legendary figure G. H. Hardy. It was clear even to an undergraduate in those days that number theory in the style of Hardy and Ramanujan was old-fashioned and did not have a great and glorious future ahead of it. Indeed Hardy in a published lecture [4] on the τ -function of Ramanujan had himself described this subject as "one of the backwaters of mathematics." The τ -function is defined as the coefficient in the modular form

$$(1) \quad \sum_{n=1}^{\infty} \tau(n)x^n = \eta^{24}(x) = x \prod_{m=1}^{\infty} (1 - x^m)^{24},$$

where $\eta(x)$ is the Dedekind eta-function. Ramanujan [5] discovered a number of remarkable arithmetical properties of $\tau(n)$. The proof and generalization of these properties by Mordell [6], Hecke [7] and others [8] played a significant part in the development of the theory of modular forms [9]. But the τ -function itself has remained a backwater, far from the mainstream of mathematics, where amateurs can dabble to their hearts' content undisturbed by competition from the professionals. Long after I became a physicist, I retained a sentimental attachment to the τ -function, and as a relief from the serious business of physics I would from time to time go back to Ramanujan's papers and meditate on the many intriguing problems that he left unsolved. Four years ago, during one of these holidays from physics, I found a new formula for the τ -function, so elegant that it is rather surprising that Ramanujan did not think of it himself. The formula is

$$(2) \quad \tau(n) = \sum \frac{(a-b)(a-c)(a-d)(a-e)(b-c)(b-d)(b-e)(c-d)(c-e)(d-e)}{1! 2! 3! 4!},$$

summed over all sets of integers a, b, c, d, e , with

$$a, b, c, d, e \equiv 1, 2, 3, 4, 5 \pmod{5},$$

$$a + b + c + d + e = 0,$$

$$a^2 + b^2 + c^2 + d^2 + e^2 = 10n.$$

This can also be written as a formula for the 24th power of the Dedekind eta-function according to (1). I was led to it by a letter from Winqvist [10], who discovered a similar formula for the 10th power of η . Winqvist also happens to be a physicist who dabbles in old-fashioned number theory in his spare time.

Pursuing these identities further by my pedestrian methods, I found that there exists a formula of the same degree of elegance as (2) for the d th power of η whenever d belongs to the following sequence of integers:

$$(3) \quad d = 3, 8, 10, 14, 15, 21, 24, 26, 28, 35, 36, \dots$$

In fact the case $d = 3$ was discovered by Jacobi [11], the case $d = 8$ by Klein and Fricke [12], and the cases $d = 14, 26$ by Atkin [13]. There I stopped. I stared for a little while at this queer list of numbers (3). As I was, for the time being, a number theorist, they made no sense to me. My mind was so well compartmentalized that I did not remember that I had met these same numbers many times in my life as a physicist. If the numbers had appeared in the context of a problem in physics, I would certainly have recognized them as the dimensions of the finite-dimensional simple Lie algebras. Except for 26. Why 26 is there I still do not know. The others all correspond to simple algebras: $A_1, A_2, B_2, G_2, A_3, B_3, A_4, D_4, A_5, B_4$ and so on. For example, $d = 24$ corresponds to the algebra A_4 , and in the structure of the formula (2) you can see the root-system of A_4 . So I missed the opportunity of discovering a deeper connection between modular forms and Lie algebras, just because the number theorist Dyson and the physicist Dyson were not speaking to each other.

This story has a happy ending. Unknown to me the English geometer, Ian MacDonald, had discovered these same formulae as special cases of a much more general theory [3]. In his theory the Lie algebras were incorporated from the beginning, and it was the connection with modular forms which came as a surprise. Anyhow, MacDonald established the connection and so picked up the opportunity which I missed. It happened also that MacDonald was at the Institute for Advanced Study in Princeton while we were both working on the problem. Since we had daughters in the same class at school, we saw each other from time to time during his year in Princeton. But since he was a mathematician and I was a physicist, we did not discuss our work. The fact that we had been thinking about the same problem while sitting so close to one another only emerged after he had gone back to Oxford. This was another missed opportunity,

but not a tragic one, since MacDonald cleaned up the whole subject very happily without any help from me. The only thing he did not clean up is the case $d = 26$, which remains a tantalizing mystery [14].

3. The Maxwell equations. I have digressed too long on this story of modular forms and Lie algebras, which I mentioned only because it is an example of a missed opportunity that happened to me personally. I now return to my main theme, the discussion of missed opportunities that were important in the historical development of mathematics.

The first clear sign of a breakdown in communication between physics and mathematics was the extraordinary lack of interest among mathematicians in James Clerk Maxwell's discovery of the laws of electromagnetism. Maxwell discovered his equations, which describe the behavior of electric and magnetic fields under the most general conditions, in the year 1861, and published a clear and definitive statement of them [15] in 1865. This was the great event of nineteenth century physics, achieving for electricity and magnetism what Newton had achieved for gravitation two hundred years earlier. Maxwell's equations contained, among other things, the explanation of light as an electromagnetic phenomenon, and the basic principles of electric power transmission and radio technology. These aspects of the theory were of primary interest to physicists and engineers. But in addition to their physical applications, Maxwell's equations had abstract mathematical qualities which were profoundly new and important. Maxwell's theory was formulated in terms of a new style of mathematical concept, a tensor field extending throughout space and time and obeying coupled partial differential equations of peculiar symmetry.

After Newton's laws of gravitational dynamics had been promulgated in 1687, the mathematicians of the eighteenth century seized hold of these laws and generalized them into the powerful mathematical theory of analytical mechanics. Through the work of Euler, Lagrange and Hamilton, the equations of Newton were analyzed and understood in depth. Out of this deep exploration of Newtonian physics, new branches of pure mathematics ultimately emerged. Lagrange distilled from the extremal properties of dynamical integrals the general principles of the calculus of variations. Fifty years later the work of Euler on geodesic motions led Gauss to the creation of differential geometry. Another fifty years later, the generalization of the Hamilton-Jacobi formulation of dynamics led Lie to the invention of Lie groups. And finally, the last gift of Newtonian physics to pure mathematics was the work of Poincaré on the qualitative behavior of orbits which led to the birth of modern topology.

But the mathematicians of the nineteenth century failed miserably to grasp the equally great opportunity offered to them in 1865 by Maxwell. If they had taken Maxwell's equations to heart as Euler took Newton's, they would have discovered, among other things, Einstein's theory of special relativity, the theory of topological groups and their linear representations, and probably large pieces of the theory of hyperbolic differential equations and functional analysis. A great part of twentieth century physics and mathematics could have been created in the nineteenth century, simply by exploring to the end the mathematical concepts to which Maxwell's equations naturally lead.

There is plenty of documentary evidence showing how Maxwell's theory appeared to the mathematicians of his time. I shall quote two short extracts to illustrate contrasting ways in which the theory was brought to their attention. Both are taken from Presidential Addresses to meetings of the British Association, then as now the chief organization in Britain dedicated to promoting the unity of science. First comes Maxwell himself, speaking in 1870, his announced topic being the relation between mathematics and physics [16].

"According to a theory of electricity which is making great progress in Germany, two electrical particles act on one another directly at a distance, but with a force which, according to Weber, depends on their relative velocity, and according to a theory hinted at by Gauss, and developed by Riemann, Lorenz, and Neumann, acts not instantaneously, but after a time depending on the distance. The power with which this theory, in the hands of these eminent men, explains every kind of electrical phenomena must be studied in order to be appreciated. Another theory of electricity which I prefer denies action at a distance and attributes electric action to tensions and pressures in an all-pervading medium, these stresses being the same in kind with those familiar to engineers, and the medium being identical with that in which light is supposed to be propagated."

It is difficult to read Maxwell's address without being infuriated by his excessive modesty, which led him to refer to his epoch-making discovery of nine years earlier as only "Another theory of electricity which I prefer." How different is his style from that of Newton, who wrote at the beginning of the third book of his *Principia* [17], "It remains that, from the same principles, I now demonstrate the frame of the System of the World." Since Maxwell himself seemed so half-hearted, it is not surprising that he did not inspire the mathematicians to throw aside their fashionable covariants and quantics and study his equations.

My second quotation is from the Oxford mathematician, Henry Smith, speaking three years later to the same audience [18]. A few months

before he spoke in 1873, Maxwell's great book on electricity and magnetism had appeared [19].

"In the course of the present year a treatise on electricity has been published by Professor Maxwell, giving a complete account of the mathematical theory of that science, No mathematician can turn over the pages of these volumes without very speedily convincing himself that they contain the first outlines (and something more than the first outlines) of a theory which has already added largely to the methods and resources of pure mathematics, and which may one day render to that abstract science services no less than those which it owes to astronomy. For electricity now, like astronomy of old, has placed before the mathematician an entirely new set of questions, requiring the creation of entirely new methods for their solution, It must be considered fortunate for the mathematicians that such a vast field of research in the application of mathematics to physical enquiries should be thrown open to them, at the very time when the scientific interest in the old mathematical astronomy has for the moment flagged,"

These words show that at least one mathematician did understand the historic nature of the challenge presented to mathematics by Maxwell's work. Smith's perception is the more remarkable, considering that he was a very pure mathematician whose best-known work is in analytic number theory. Unfortunately he was then 46 years old, too old to be a pioneer in a new field. No doubt he contented himself with "turning over the pages of these volumes" and then returned with relief to his familiar ternary forms. The young men, who might have been stimulated by his words to create new fields of mathematics, were not listening. Hermann Minkowski and Jacques Hadamard, who were to achieve in the twentieth century the fulfilment of some of Smith's prophecies, were boys of nine and eight at the time he spoke. Élie Cartan was three; Hermann Weyl, Jean Leray and Harish-Chandra were not yet born.

Hermann Minkowski had something to say thirty-five years later about the opportunity which the mathematicians had missed. He was talking to the German equivalent of the British Association, three years after Einstein's discovery of special relativity. He pointed out that the mathematical basis of Einstein's discovery lies in an incompatibility between two groups of transformations of space and time coordinates. On the one hand, the equations of Newtonian mechanics are invariant under a group G_∞ which physicists now call the *Galilei group*. The group G_∞ is six-dimensional; it is generated by three rotations of the space coordinates, and three uniform velocity transformations of the form

$$(4) \quad x \rightarrow x - ut, \quad t \rightarrow t.$$

On the other hand, the Maxwell equations in the absence of matter are invariant under a group G_c which physicists call the *Lorentz group*. The group G_c is also six-dimensional; it is generated by three rotations as before, together with three Lorentz transformations of the form

$$(5) \quad \begin{aligned} x &\rightarrow \beta(x - ut), & t &\rightarrow \beta(t - (ux/c^2)), \\ \beta &= (1 - (u/c)^2)^{-1/2}, \end{aligned}$$

where c is the velocity of light. From a purely mathematical point of view, G_c has a simpler structure than G_∞ . In fact G_c is a real noncompact form of the semisimple Lie algebra $A_1 \times A_1$, whereas G_∞ is not semisimple. I quote now from Minkowski's 1908 lecture [2].

"Group G_c in the limit when $c = \infty$, that is the group G_∞ , becomes no other than that complete group which is appropriate to Newtonian mechanics. This being so, and since G_c is mathematically more intelligible than G_∞ , it looks as though the thought might have struck some mathematician, fancy-free, that after all, as a matter of fact, natural phenomena do not possess an invariance with the group G_∞ , but rather with the group G_c , c being finite and determinate, but in ordinary units of measure, extremely great. Such a premonition would have been an extraordinary triumph for pure mathematics. Well, mathematics, though it now can display only staircase-wit, has the satisfaction of being wise after the event, and is able, thanks to its happy antecedents, with its senses sharpened by an unhampered outlook to far horizons, to grasp forthwith the far-reaching consequences of such a metamorphosis of our concept of nature."

Why were the mathematicians of the later nineteenth century blind to these possibilities which Smith had so clearly foreshadowed? There are many reasons. If Maxwell had written in a style as lucid and as confident as that of Newton, the mathematicians would have been more inclined to take him seriously. Another reason for the mathematicians' indifference was the fact that Maxwell's equations were not generally accepted even by physicists for twenty years after their discovery. Until Hertz demonstrated the existence of radio waves in 1885, the majority of physicists considered Maxwell's theory to be a speculative hypothesis [20]. Mathematicians who had themselves lost touch with physics were not able to make an independent assessment of the theory's merits. Lastly, and perhaps most importantly, the mathematicians ignored Maxwell because mathematics in the later nineteenth century had developed in quite other directions. Mathematicians were busy with the theory of functions of complex variables, with analytic number theory, with algebraic forms and invariants. The flowering of these subjects had given the mathematicians

definite tastes and aesthetic standards, into which the new physics of Maxwell would not easily fit.

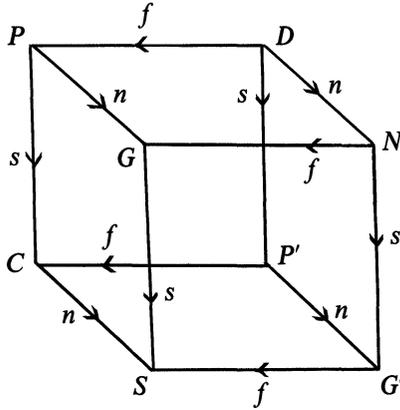
4. Kinematical groups. The story of the Maxwell equations has a postscript, in which the pure mathematicians again missed an opportunity, though not one of such major importance as that which they missed in 1873. Nobody noticed that Minkowski in his 1908 lecture failed to carry his argument to its logical conclusion. Minkowski did not mention the fact that the Maxwell equations are invariant under the trivial Abelian group T_4 of translations of the space-time coordinates. The natural invariance group of the Maxwell theory is not the six-dimensional Lorentz group G_c but the ten-dimensional Poincaré group P which is a semi direct product of G_c with T_4 . Similarly, the invariance group of Newtonian mechanics is not the six-dimensional G_∞ but the ten-dimensional Galilei group G which is a semidirect product of G_∞ with T_4 . Neither P nor G is a semisimple group.

With hindsight it is easy to see that Minkowski's logic ought to have given somebody the idea that there exists a simple group D of which the nonsemisimple group P is a degenerate limit, in exactly the same way as the semisimple G_c has the nonsemisimple G_∞ as a degenerate limit. This D is the DeSitter group, a real noncompact form of the simple Lie algebra B_2 . Following Minkowski's argument, a pure mathematician might easily have conjectured in 1908 that the true invariance group of the universe should be D rather than P . D is in fact the invariance group of an empty expanding universe whose radius of curvature R is a linear function of time. D degenerates to P in the flat-space limit $R \rightarrow \infty$, just as G_c degenerates to G_∞ in the Newtonian limit $c \rightarrow \infty$.

Suppose that somebody had been bold enough in 1908 to take this idea seriously. He would have correctly predicted the expansion of the universe twenty years before it was discovered observationally by Hubble. More importantly, he would have been led to postulate the curvature of space-time, and so he would have considerably eased the path which led to general relativity. Luckily, Einstein was able to reach general relativity the hard way, without having his path eased for him by anybody. DeSitter in fact discovered his model of an expanding universe a year after he learned of Einstein's theory.

Only after an interval of sixty years has the logical framework of Minkowski's argument been completed. In a beautiful analysis published in 1968, Bacry and Lévy-Leblond [21] proved that, subject to a certain reasonable set of consistency conditions, there exist precisely eight kinematical groups. By a kinematical group they mean a group which can serve, consistently with the general principles of quantum mechanics, as

the invariance group of a universe possessing properties of uniformity and isotropy. Of the eight groups, only D is simple, and the other seven are derived from it by applying three limiting processes in all possible combinations. Specifically, let f denote the flat-space limit $R \rightarrow \infty$, let n denote the Newtonian limit $c \rightarrow \infty$, and let s denote the static limit $c \rightarrow 0$. The eight groups can then be visualized as the vertices of a cube.



D , $P = fD$ and $G = nfD$ are the only kinematical groups that correspond to orthodox physical universes. But the other five groups are just as good, mathematically speaking. The most interesting of the heterodox groups are $N = nD$ and $C = sfD$. N describes a Newtonian universe with curved space-time. C describes a universe in which space is absolute, in contrast to the Galilei group G which has time absolute. The group C was discovered by Lévy-Leblond [22] and called by him the Carroll group. In the Carroll universe all objects have zero velocity although they may have nonzero momentum. Carroll was a pure mathematician who had already foreseen this possibility in 1871 [23]:

“A slow sort of country,” said the Queen, “Now, here, you see, it takes all the running you can do, to keep in the same place.”

But his mathematical colleagues once again missed an opportunity by failing to take him seriously.

5. Quaternions and vectors. It is appropriate in a Gibbs lecture to mention another missed opportunity which Gibbs himself presented to the mathematicians of his time. The opportunity had existed for forty years before Gibbs [24] explicitly called attention to it in 1886. But another forty years were to pass before the mathematicians fully responded to it.

In the year 1844 two remarkable events occurred, the publication by Hamilton [25] of his discovery of quaternions, and the publication by Grassmann [26] of his "Ausdehnungslehre." With the advantage of hindsight we can see that Grassmann's was the greater contribution to mathematics, containing the germ of many of the concepts of modern algebra, and including vector analysis as a special case. However, Grassmann was an obscure high-school teacher in Stettin, while Hamilton was the world-famous mathematician whose official titles occupy six lines of print after his name at the beginning of his 1844 paper. So it is regrettable, but not surprising, that quaternions were hailed as a great discovery, while Grassmann had to wait 23 years before his work received any recognition at all from professional mathematicians. When Grassmann's work finally became known, mathematicians were divided into quaternionists and antiquaternionists, and were spending more energy in polemical arguments for and against quaternions than in trying to understand how Grassmann and Hamilton might be fitted together into a larger scheme of things. So it was left to the physicist Gibbs to present for the first time in his 1886 lecture the essential ideas of Grassmann and Hamilton side by side. The last words of his lecture are, "We begin by studying multiple algebras; we end, I think, by studying MULTIPLE ALGEBRA."

I do not know how many pure mathematicians heard or read Gibbs' lecture. If they had studied it carefully, they would soon have noticed that Gibbs had not really succeeded in unifying the notions of quaternion and vector. On the contrary, by putting the two notions side by side he had made explicit the lack of any real compatibility between them. His lecture ought to have suggested to any attentive mathematician the question, "How can it happen that the properties of three-dimensional space are represented equally well by two quite different and incompatible algebraic structures?" If this question had once been clearly asked, the answer would almost certainly have been forthcoming. And the answer would have led inevitably to a complete theory of the single-valued and double-valued representations of the three-dimensional rotation group. The vectors are the simplest nontrivial single-valued representation, and the quaternions are the simplest double-valued representation. Also, the quaternions are the prototype of what later were called spinor representations. The development of the theory of spinor representations, which was actually begun by Elie Cartan in 1913 and completed during the 1930's [27] with substantial help from the physicists Pauli and Dirac, might have been accelerated by approximately 40 years. It is impossible to say what effects such an accelerated development would have had on other branches of pure mathematics, but the effects could hardly have failed to be substantial.

6. General coordinate invariance. Up to now, my examples of missed opportunities have been mathematical discoveries which actually occurred, although they could have occurred a long time earlier. In such cases one can be sure that an opportunity existed, but it existed only in the past. I now come to the more difficult task of identifying missed opportunities that are still open. Here one can no longer be sure that the opportunity is real, but if it is real then it has the virtue of existing in the present.

The past opportunities which I discussed have one important feature in common. In every case there was an empirical finding that two disparate or incompatible mathematical concepts were juxtaposed in the description of a single situation. Taking the four examples in turn, the pairs of disparate concepts were respectively: modular functions and Lie algebras, field equations and particle dynamics, Lorentz invariance and Galilean invariance, quaternion algebra and Grassmann algebra. In each case the opportunity offered to the pure mathematician was to create a wider conceptual framework within which the pair of disparate elements would find a harmonious coexistence. I take this to be my methodological principle in looking for opportunities that are still open. I look for situations in which the juxtaposition of a pair of incompatible concepts is acknowledged but unexplained.

The most glaring incompatibility of concepts in contemporary physics is that between Einstein's principle of general coordinate invariance and all the modern schemes for a quantum-mechanical description of nature. Einstein based his theory of general relativity [28] on the principle that God did not attach any preferred labels to the points of space-time. This principle requires that the laws of physics should be invariant under the Einstein group E , which consists of all one-to-one and twice-differentiable transformations of the coordinates. By making full use of the invariance under E , Einstein was able to deduce the precise form of his law of gravitation from general requirements of mathematical simplicity without any arbitrariness. He was also able to reformulate the whole of classical physics (electromagnetism and hydrodynamics) in E -invariant fashion, and so determine unambiguously the mutual interactions of matter, radiation and gravitation within the classical domain. There is no part of physics more coherent mathematically and more satisfying aesthetically than this classical theory of Einstein based upon E -invariance.

On the other hand, all the currently viable formalisms for describing nature quantum-mechanically use a much smaller invariance group. The analysis of Bacry and Lévy-Leblond [21] indicates the extreme range of quantum-mechanical kinematical groups that have been contemplated. In practice all serious quantum-mechanical theories are based either on the Poincaré group P or the Galilei group G . This means that a class of

preferred inertial coordinate-systems is postulated a priori, in flat contradiction to Einstein's principle. The contradiction is particularly uncomfortable, because Einstein's principle of general coordinate invariance has such an attractive quality of absoluteness. A physicist's intuition tells him that, if Einstein's principle is valid at all, it ought to be valid for the whole of physics, quantum-mechanical as well as classical. If the principle were not universally valid, it is difficult to understand why Einstein achieved such deeply coherent insights into nature by assuming it to be so.

To make the mathematical incompatibility more definite, I will focus attention on one of the competing schemes for describing a quantum-mechanical universe. I choose the scheme which is most carefully based on rigorous mathematical definitions and which is also general enough to encompass a wide variety of physical systems. This scheme is the "Algebra of Local Observables" of Haag and Kastler [29]. The six axioms of Haag and Kastler are the following.

(1) *Existence of local observables.* To every region B (i.e. open set with compact closure) in 4-dimensional space-time there corresponds an abstract C^* -algebra $\mathcal{A}(B)$.

(2) *Isotony.* If $B_1 \supset B_2$ then $\mathcal{A}(B_1) \supset \mathcal{A}(B_2)$, and $\mathcal{A}(B_1)$ and $\mathcal{A}(B_2)$ have a common unit element.

(3) *Existence of quasilocal observables.* The union of all $\mathcal{A}(B)$ is a normed $*$ -algebra, the completion of which is a C^* -algebra \mathcal{A} , the algebra of quasilocal observables. It is supposed that all physically measurable quantities are elements of \mathcal{A} .

(4) *Poincaré invariance.* To every element L of the Poincaré group there corresponds an automorphism α_L of \mathcal{A} , such that $\alpha_L(\mathcal{A}(B)) = \mathcal{A}(LB)$ for every region B .

(5) *Local commutativity.* If two regions B_1 and B_2 are completely spacelike with respect to each other (i.e. if B_2 lies outside every light-cone with vertex in B_1), then $\mathcal{A}(B_1)$ and $\mathcal{A}(B_2)$ commute.

(6) *Primitivity.* \mathcal{A} admits a faithful algebraically irreducible representation.

These axioms, taken together with the axioms defining a C^* -algebra [30], are a distillation into abstract mathematical language of all the general truths that we have learned about the physics of microscopic systems during the last 50 years. They describe a mathematical structure of great elegance whose properties correspond in many respects to the facts of experimental physics. In some sense, the axioms represent the most serious attempt that has yet been made to define precisely what physicists mean by the words "observability, causality, locality, relativistic invariance," which they are constantly using or abusing in their everyday speech.

If we look at the axioms in detail, we see that (1), (2), (3) and (6) are consistent with Einstein's general coordinate invariance, but (4) and (5) are inconsistent with it. Axioms (4) and (5), the axioms of Poincaré invariance and local commutativity, require the Poincaré group to be built into the structure of space-time. If we try to replace the Poincaré group P by the Einstein group E , we have no way to define a space-like relationship between two regions, and axiom (5) becomes meaningless. I therefore propose as an outstanding opportunity still open to the pure mathematicians, *to create a mathematical structure preserving the main features of the Haag-Kastler axioms but possessing E -invariance instead of P -invariance.*

I had better warn any mathematician who intends to respond to my challenge that his task will not be easy. No merely formal rearrangement of the Haag-Kastler axioms can possibly be sufficient. For we know that Einstein could construct his E -invariant classical theory of 1916 only by bringing in the full resources of Riemannian differential geometry. He needed a metric tensor to give his space-time a structure independent of coordinate-systems. Therefore an E -invariant axiom of local commutativity to replace axiom (5) will require at least some quantum-mechanical analog of Riemannian geometry. Some analog of a metric tensor must be introduced in order to give a meaning to space-like separation. The answer to my challenge will necessarily involve a delicate weaving together of concepts from differential geometry, functional analysis, and abstract algebra. With these words of warning I leave the problem to you.

7. Feynman sums. My second example of a still open opportunity also concerns very basic aspects of quantum physics. Twenty years ago, before C^* -algebras became fashionable, Richard Feynman [31] gave a description of relativistic quantum field theory in terms of a naive physical picture which he called "sum over histories." His description seems to make sense as a qualitative guide to the understanding of physical processes, but it makes no sense at all as a mathematical definition. Mathematical rigor is the last thing that Feynman was ever concerned about.

Feynman's picture contains the following elements. A *history* H is defined by specifying the value $\varphi_H(x)$ of a classical vector-valued function φ at every point x of space-time. A *space-like surface* σ is a surface separating space-time into two parts (past and future) such that every two points on σ are separated by a space-like interval. A *state* ψ_σ on σ is a complex-valued functional of the function $\varphi(x)$ with x restricted to σ . The states on σ form a complex linear manifold D_σ . The dynamics of the universe are determined by a *Lagrangian* $L(\varphi)$ which is a real-valued

function of the value and derivatives of φ at a single point x . A *classical local observable* A_B is a complex-valued functional of the function $\varphi(x)$ with x restricted to a bounded region B . A *quantum local observable* $O(A_B)$ is defined as a complex-valued bilinear form on D_σ and D_τ , where σ and τ are space-like surfaces such that B lies wholly in the future of σ and in the past of τ . The classical A_B determines the quantum $O(A_B)$ according to the following rule:

$$(6) \quad (\chi_\tau, O(A_B)\psi_\sigma) = \sum_H (\chi_\tau(\varphi_H))^* A_B(\varphi_H)\psi_\sigma(\varphi_H) \exp \left[i \int_\sigma^\tau L(\varphi_H(x)) d_4x \right]$$

The \sum_H here denotes the famous "sum over histories," which is supposed to include all possible classical functions $\varphi_H(x)$ defined on the 4-dimensional slice of space-time between σ and τ . Taken literally, this \sum_H is mathematical nonsense. The essential problem is to find a mathematical framework within which the definition (6) can be given a precise meaning.

In this problem the juxtaposition of a pair of incompatible mathematical elements occurs in the following way. On the one hand, the formula (6), if manipulated in a purely formal style without any regard for rigorous justification, gives all the right answers. It gives, for example, the right Lagrangian field equations satisfied by the quantum observable $O(\varphi(x))$, the right commutation relations between the observables $O(A_B)$ and the generators of the Poincaré group, the right commutation relation of Peierls [32] between two local observables, and the right definition of the scattering matrix in the limit when the surfaces σ, τ tend to the infinite past and the infinite future respectively. On the other hand, the existing theories of normed vector spaces impose drastic requirements which Feynman's functional manifolds seem to have no hope of satisfying. We have a juxtaposition of a kind which has frequently occurred in formative stages of the history of mathematics. We have a method which leads to correct conclusions, and a rigorous theory which forbids the use of the method.

According to (6), the identity operator I is associated with the special bilinear form

$$(7) \quad (\chi_\tau, \psi_\sigma) = \sum_H (\chi_\tau(\varphi_H))^* \psi_\sigma(\varphi_H) \exp \left[i \int_\sigma^\tau L(\varphi_H(x)) d_4x \right],$$

which maps D_τ onto the dual space D'_σ of D_σ . Call this mapping $I_{\tau\sigma}$, and suppose (of course without proof) that it has an inverse $J_{\sigma\tau}$ mapping D'_σ onto D_τ . The quantum observable $O(A_B)$ is also a bilinear form mapping D_τ onto D'_σ . Therefore the product

$$Q(A_B) = J_{\sigma\tau} O(A_B)$$

is a linear mapping of D_t onto itself, i.e. a linear operator on the manifold of states. So we have constructed a linear operator $Q(A_B)$ to represent each quantum local observable. If this construction could be rigorously justified, then we could define a norm for the bounded $Q(A_B)$ and so generate a C^* -algebra $\mathcal{A}(B)$ from the bounded local observables associated with the region B . If we had started with a Lagrangian $L(\varphi)$ invariant under the Poincaré group, we could expect that our C^* -algebras of local observables would satisfy the Haag-Kastler axioms. We would then have achieved a completely constructive definition of a relativistic quantum field theory.

This program of constructing a P -invariant quantum field theory by means of Feynman sums can actually be carried through in the case when L is a quadratic function of φ and its first derivatives. In this case the sum over histories in (6) can be given a rigorous meaning as the complex analog of a Wiener integral [33]. The theory then gives a completely satisfactory description of an assemblage of quantum-mechanical particles possessing the properties of wave-particle duality as they are observed in nature. Unfortunately one important feature of real particles is lacking. In a theory with quadratic Lagrangian the particles do not interact. The quadratic case is mathematically interesting but physically trivial.

A great deal of work has been done by physicists using models in which the Lagrangian is of the form $L = L_0 + L_1$ where L_0 is quadratic and L_1 is nonquadratic but small. In these models everything is calculated by perturbation theory using an expansion in powers of L_1 . To any finite order in L_1 the Feynman sums can be given a meaning. The models then describe interacting particles and give results which accurately reproduce many features of the real world. It is difficult to believe that these models, which in perturbation theory behave so astonishingly well, are mathematically meaningless. But all attempts to give them a rigorous mathematical definition have so far failed. The opportunity has been open to mathematicians for 20 years, and is still open to them, *to construct a conceptual scheme which will legalize the use of Feynman sums such as (6) with suitable Lagrangians which are not quadratic*. Whoever achieves this will have created the first rigorous theory of quantized relativistic particles with local interactions in a 4-dimensional space-time.

So far I have discussed Feynman sums only in connection with theories invariant under the Poincaré group. But unlike the Haag-Kastler axioms, Feynman sums can be defined independently of the Poincaré group. In particular, if we choose a Lagrangian which is formally an invariant scalar-density under general coordinate transformations, then the Feynman sums ought to give rise to an E -invariant quantum field theory.

Admittedly, E -invariant Lagrangians, such as the Lagrangian in Einstein's theory of gravitation, are apt to be awkward and pathological in various ways. It may well happen that, even after a way has been found to define rigorously the Feynman sums in nontrivial P -invariant theories, the E -invariant sums still defy interpretation. I therefore state as a separate challenge to the mathematicians that they should try to achieve a *rigorous definition of Feynman sums which are invariant under general coordinate transformations*. A successful solution of this problem would kill two birds with one stone; it would simultaneously demonstrate that the axioms of quantum field theory have a nontrivial realization, and that these axioms can be reconciled with the principles of general relativity.

8. Conclusion. There are several other still open opportunities which I could discuss at length if I had more time. One is to incorporate the economic minimax theorem of von Neumann [34] into a coherent analytical structure, in the same sense in which Hilbert incorporated the classical eigenvalue problems of physics into a general theory of linear operators. Another opportunity exists in the theory of resonances between planetary orbits, which was the subject of a Gibbs lecture 44 years ago [35]. To understand the existence and stability of such resonances is a problem of Newtonian mechanics which has been begging for solution since the time of Laplace. In spite of heroic efforts by such mathematicians as Poincaré and Jürgen Moser, the resonances remain a mystery. It came as a surprise to the theoreticians when a numerical analysis [36] recently revealed that the orbits of Neptune and Pluto are locked in a stable resonance.

Undoubtedly there exist many more missed opportunities to create new branches of pure mathematics out of old problems of applied science. But I think I have mentioned enough of them to demonstrate that there is some truth in the words of Hadamard [37] with which I began this lecture. Hadamard, incidentally, practiced what he preached.²

REFERENCES

1. D. Hilbert, *Mathematische Probleme*, Lecture to the Second Internat. Congress of Math. (Paris, 1900), Arch. Math. und Phys. (3) 1 (1901), 44–63; 213–237; English transl., Bull. Amer. Math. Soc. 8 (1902), 437–479.

² NOTE ADDED IN PROOF. Several people who heard this lecture have told me that the τ -function is not such a backwater as I had supposed. The τ -function appears in a modern context in B. J. Birch, *How the number of points of an elliptic curve over a fixed prime field varies*, J. London Math. Soc. 43 (1968), 57–60, MR 37 # 6242, and in P. Deligne, *Formes modulaires et représentations l -adiques*, Séminaire Bourbaki, 1968/69, No. 355. A new attack on the problem of defining Feynman sums in general relativity is described in L. D. Faddeev, *Symplectic structure and quantization of the Einstein gravitation theory*, Proc. Internat. Math. Congress, Nice, 1970, E2 (157), 78–82.

2. H. Minkowski, *Raum und Zeit*, Lecture to the 80th Assembly of Natural Scientists (Köln, 1908), Phys. Z. **10** (1909), 104–111. English transl., *The principle of Relativity*, Aberdeen Univ. Press, Aberdeen, 1923.
3. I. G. MacDonald, *Affine root systems and Dedekind's η -function*, Invent. Math. **15** (1972), 91–143. See also R. V. Moody, *Euclidean Lie algebras*, Canad. J. Math. **21** (1969), 1432–1454. MR **41** # 287.
4. G. H. Hardy, *Ramanujan*, Cambridge Univ. Press, Cambridge; Macmillan, New York, 1940, Chap. 10, p. 161. MR **3**,71.
5. S. Ramanujan, *On certain arithmetical functions*, Trans. Cambridge Philos. Soc. **22** (1916), 159–184.
6. L. J. Mordell, *On Mr. Ramanujan's empirical expansions of modular functions*, Proc. Cambridge Philos. Soc. **19** (1917), 117–124.
7. E. Hecke, *Über Modulfunktionen und die Dirichletschen Reihen mit Eulerscher Produktentwicklung*, Math. Ann. **114** (1937), 1–28; 316–351.
8. M. Newman, *An identity for the coefficients of certain modular forms*, J. London Math. Soc. **30** (1955), 488–493. MR **17**,15; A. O. L. Atkin, *Ramanujan congruences for p - k (n)*, Canad. J. Math. **20** (1968), 67–78; corrigendum, *ibid.* **21** (1969), 256. MR **38** # 2098.
9. R. C. Gunning, *Lectures on modular forms*, Ann. of Math. Studies, no. 48, Princeton Univ. Press, Princeton, N.J., 1962. MR **24** # A2664; A. Ogg, *Modular forms and Dirichlet series*, Benjamin, New York, 1969. MR **42** # 1648.
10. L. Winquist, *An elementary proof of $p(11m + 6) \equiv 0 \pmod{11}$* , J. Combinatorial Theory **6** (1969), 56–59. MR **38** # 4434. Winquist sent this proof to me in January 1968.
11. C.G.J. Jacobi, *Fundamenta nova theoriae functionum ellipticarum*, Königsberg, 1829, 66, Eq. (5).
12. F. Klein and R. Fricke, *Vorlesungen über die Theorie der elliptischen Modulfunktionen*. Vol. 2, Teubner, Leipzig, 1892, p. 373.
13. A.O.L. Atkin had known of these formulae for some time before 1968 when he sent me proofs of them. His work on them is still unpublished.
14. The following special case of Atkin's formula for $d = 26$ brings out clearly the analogy with the formula (2) for $\tau(n)$. Write

$$\eta^{26}(x) = \sum \tau_{26}(n)x^{n/12}, \quad n \equiv 1 \pmod{12}.$$

For n prime we have unique representations

$$n = a^2 + b^2 = c^2 + 3d^2, \quad a \equiv 0 \pmod{3},$$

and

$$\begin{aligned} \tau_{26}(n) = & (-1)^a (2^3 \cdot 3^4 \cdot 11^2 \cdot 13)^{-1} (a-c)(a+c)(b-c)(b+c) \\ & \cdot (2a+c+3d)(2a+c-3d)(2a-c+3d)(2a-c-3d) \\ & \cdot (2b+c+3d)(2b+c-3d)(2b-c+3d)(2b-c-3d). \end{aligned}$$

But the factors here do not correspond in any obvious way to the system of roots of a Lie algebra.

15. J. Clerk Maxwell, *A dynamical theory of the electromagnetic field*, Philos. Trans. Roy. Soc. (London) **155** (1865), 459–512.

16. ———, Presidential Address to Section A (Mathematical and Physical Sciences) of the British Association, Liverpool, 1870. Nature **2** (1870), 419–422.

17. I. Newton, *Mathematical principles of natural philosophy*, translated into English by Andrew Motte in 1729, edited by F. Cajori, Univ. of California Press, Berkeley, Calif., 1946, p. 397.

18. Henry J.S. Smith, Presidential Address to Section A (Mathematical and Physical Sciences) of the British Association, Bradford, 1873. Nature **8** (1873), 448–452.

19. J. Clerk Maxwell, *A treatise on electricity and magnetism*, Oxford Univ. Press, Oxford, 1873.

20. The physicist Michael Pupin, in his autobiography, *From immigrant to inventor*, Charles Scribner's Sons, 1924, gives a vivid account of the difficulties encountered by a student trying to learn the Maxwell theory in the 1880's. Pupin traveled from America to England searching in vain for somebody who had mastered the theory; he finally succeeded in learning it from Helmholtz in Berlin.

21. H. Bacry and J.-M. Lévy-Leblond, *Possible kinematics*, J. Mathematical Phys. **9** (1968), 1605–1614. MR **38** # 6821. To save time I have slightly misstated their conclusion; each of the groups D , P and N can occur in two alternative forms, so that the number of possibilities is strictly speaking 11 rather than 8.

22. J.-M. Lévy-Leblond, *Une nouvelle limite non-relativiste du groupe de Poincaré*, Ann. Inst. H. Poincaré Sect. A **3** (1965), 1–12. MR **33** # 1125.

23. L. Carroll, *Through the looking-glass, and what Alice found there*, Macmillan, London, 1871.
24. J. Willard Gibbs, *On multiple algebra*, Vice-Presidential Address to the Section of Mathematics and Astronomy of the American Association for the Advancement of Science, Proc. Amer. Assoc. Adv. Sci. **35** (1886), 37–66.
25. W. R. Hamilton, *On quaternions; or on a new system of imaginaries in algebra*, Philos. Mag. **25** (1844), 10–13.
26. H. Grassmann, *Die lineale Ausdehnungslehre, ein neuer Zweig der Mathematik*, Otto Wigand, Leipzig, 1844.
27. See R. Brauer and H. Weyl, *Spinors in n dimensions*, Amer. J. Math. **57** (1935), 425–449.
28. A. Einstein, *Die Grundlage der allgemeinen Relativitätstheorie*, Ann. Phys. **49** (1916), 769–822.
29. R. Haag and D. Kastler, *An algebraic approach to quantum field theory*, J. Mathematical Phys. **5** (1964), 848–861. MR **29** # 3144.
30. J. Dixmier, *Les C^* -algèbres et leurs représentations*, Cahiers Scientifiques, fasc. 29, Gauthier-Villars, Paris, 1964. MR **30** # 1404.
31. R. P. Feynman, *Mathematical formulation of the quantum theory of electromagnetic interaction*, Phys. Rev. (2) **80** (1950), 440–457. MR **12**,889; R. P. Feynman, *Space-time approach to non-relativistic quantum mechanics*, Rev. Modern Phys. **20** (1948), 367–387. MR **10**,224.
32. R. E. Peierls, *The commutation laws of relativistic field theory*, Proc. Roy. Soc. London Ser. A **214** (1952), 143–157. MR **14**,520.
33. I. M. Gelfand and A. M. Jaglom, *Integration in functional spaces and its applications in quantum physics*, Uspehi Mat. Nauk **11** (1956), no. 1 (67), 77–114; English transl., J. Mathematical Phys. **1** (1960), 48–69. MR **22** # 3455, and its bibliography is complete up to 1956. Substantial progress in justifying the use of Feynman sums in a nonrelativistic context has been made more recently; see E. Nelson, *Feynman integrals and the Schrödinger equation*, J. Mathematical Phys. **5** (1964), 332–343. MR **28** # 4397, and C. M. DeWitt, *Feynman's path integral, definition without limiting procedure*, Univ. of Texas, Austin, Tex., 1972 (preprint).
34. J. von Neumann, *Über ein ökonomisches Gleichungssystem und eine Verallgemeinerung des Brouwerschen Fixpunktsatzes*, Ergebnisse eines mathematischen Seminars, edited by K. Menger, Wien, 1938; English transl., Rev. Economic Studies **13** (1945), 1–9.
35. E. W. Brown, *Resonance in the solar system*, Bull. Amer. Math. Soc. **34** (1928), 265–289.
36. C. J. Cohen and E. C. Hubbard, *Libration of the close approaches of Pluto to Neptune*, Astronom. J. **70** (1965), 10–13.
37. J. Hadamard, *The psychology of invention in the mathematical field*, Princeton Univ. Press, Princeton, N.J., 1945, Chap. 4. MR **6**,198.