IC/80/39

INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS

THE NATURE OF THE "ULTIMATE" EXPLANATION IN PHYSICS

Abdus Salam

INTERNATIONAL ATOMIC ENERGY AGENCY

UNITED NATIONS EDUCATIONAL, SCIENTIFIC AND CULTURAL ORGANIZATION

1980 MIRAMARE-TRIESTE

IC/80/39

International Atomic Energy Agency
and
United Nations Educational Scientific and Cultural Organization

INTERNATIONAL CENTRE FOR THEORETICAL PHYSICS

THE NATURE OF THE "ULTIMATE" EXPLANATION IN PHYSICS *

Abdus Salam
International Centre for Theoretical Physics, Trieste, Italy,
and
Imperial College, London, England.

MIRAMARE - TRIESTE
March 1980

* Herbert Spencer Lecture, Oxford University, 12 November 1979 (to appear in "Scientific Explanation", Oxford University Press).

"Experiment alone can decide on truth" ... "But the axiomatic basis of physics cannot be extracted from experiment.".

— Einstein, Herbert Spencer Lecture, June 1933

§1. All science - physics in particular - is concerned with discovering WHY things happen as they do. The WHYS so adduced must clearly be "deeper", more universal, more axiomatic, less susceptible to direct experimental testing, than the immediate phenomena we seek to explain. And it is also well-known, that it is the WHYS of one generation which are often the points of departure for the next, to whom the earlier WHYS can appear subjective, conditioned by "unscientific" thinking, even wrong. The glory of science is that this notwithstanding, we often arrive at correct predictions - at least to the extent of the experimental accuracies achievable and often better. I wish to speak about this continuing, ever-sharpening process about the WHYS of physics in the context of the fundamental unification of physical forces on which our generation is engaged.

I can summarize my remarks in terms of three propositions:

1) The physics of the last century ascribed its deeper WHYS to an all-pervading mechanical aether. Einstein killed this aether, but he substituted for it, something terribly close in spirit - a dynamical space-time manifold. Following Einstein, the deepest WHYS of to-day's physics are to be found as manifestations of what we choose to assume as the basic attributes of the space-time manifold.

2) So far as dynamics is concerned, our final court of appeal, if all else fails, is the Bootstrap mechanism, the principle of self-consistency of the Universe. This principle may be traced back to the teleological dictum of Leibnitz - so savagely satirised by Voltaire in Candide - "The Universe is as it is for what else could it be.".

3) And finally, there are the Laws of Impotence - so named by Max Born - which all WHYS must respect. These laws of impotence - the glory of the physics of the 20th century - consist of not-to-be-questioned admonitions like: thou shalt not conceive of velocities greater than that of light, to transmit signals; thou shalt quantize angular momentum in units of the Planck's constant ($\hbar$) .

There are other requirements governing the desirable WHYS, like economy of concepts and simplicity (Occam's razor), like eschewing of over-subtlety, like beauty of the mathematics to be used (which somehow appears linked with its unreasonable efficacy). But these are well-known ideas and do not need elaboration.

§2. To illustrate my remarks, and in particular the questioning by one generation of the WHYS which led the generation before to (relative) truth, consider the classic example of the laws of planetary orbits and celestial gravity theory, associated with the names of Kepler, Newton and Einstein.

Kepler, the first man to give a quantitative description of laws of planetary motion describes thus how he was led to their discovery.

"God reflected on the difference between the curved and the straight and preferred the nobility of the curved.".

"Among bodies, omit ... the irregular ones, and only retain those whose faces are equal in side and in angle. There remain five regular bodies of the Greeks: cube, pyramid, dodecahedron, icosahedron and octahedron. ... If the five bodies be fitted into one another and if circles be described both inside and outside all of them, then we obtain precisely the number six of circles. ... Copernicus has taken just six orbits of this kind, pairs of which are precisely related by the fact that those five bodies fit most perfectly into them.".

Would this type of reasoning be considered "scientific" to-day*?

Kepler described Copernicus as a "blind man feeling his way with a staff". It must have been this act of hubris which in turn had its nemesis in Köestler's description of Kepler as a "sleep-walker".

Kepler was followed by Newton, who washed his hands of the entire search for WHY; "But hitherto I have not been able to discover the cause of ... gravity from phenomena and I frame no hypotheses ... Hypothesis ... has no place in experimental philosophy.".

On this attitude of Newton, Einstein had this to say: "We now realize with special clarity, how much in error are those theorists who believe that theory comes inductively from experiment. Even the great Newton could not free himself from this error (Hypotheses non fingo).".

But had Newton built no hypothesis into his gravity theory? According to Einstein, he had. This was the hypothesis that the gravitational charge (m) which occurs in Newton's Force Law ($F = \frac{m_1 m_2}{r^2}$) exactly equals inertial mass – the quantity of matter contained in the bodies which mutually attract. This is the so-called Equivalence Principle.

§3. To see the force of Einstein's remark about Newton's assumption of the equality of gravitational charge with inertial mass, consider a hydrogen atom which consists of a proton and an electron. In making up the atom, the electron and the proton attract each other both electrically as well as gravitationally. The inertial mass of the atom

---

* Before we dismiss Kepler's reasoning, reflect on our own generation's partiality for the eight-fold way, or for the exceptional Lie groups as candidates for symmetry groups in particle physics, stemming as this partiality usually does from the mathematical "nobility" of these particular constructs!

equals proton's mass plus electron's mass minus the electrical, as well as the gravitational, binding energies. The ratio of the summed masses of the proton and the electron to the two varieties of binding energies is of the order of $1 : 10^{-8} : 10^{-47}$. Now, Eötvös (in the nineteenth century) in his celebrated torsion experiment, had in fact demonstrated that gravitational charge does equal inertial mass to the extent that, for the hydrogen atom for example, the electrical binding energy ($10^{-8} : 1$) contributes equally to both. But what about the gravitational binding energy? Does the tiny relative number (10^{-47}) – ascribable to gravitational binding – also affect inertial mass and gravitational charge equally? What would Newton say?

Einstein's own answer was unambiguous. His WHY for the existence of the gravitational force ascribes this force to space-time dynamics, to the curvature of the four dimensional space-time. His theory incorporates a "strong equivalence" of gravitational charge with inertial mass. But there were rival theories – like those of Brans-Dicke's extension of Einstein's – which denied this equivalence so far as the gravitational binding-energy is concerned. According to these theories, a part of this relative 10^{-47} would not show up in the gravitational charge.

The issue between Einstein and Brans-Dicke was joined, in March 1976, in two beautiful experiments, independently carried out by two teams; one led by Shapiro, the other by Dicke himself. These epic experiments consisted of measuring the mean (Kepler) positions of the earth and the moon to ± 30 cms. through lunar laser ranging measurements. For heavenly test bodies as massive as these, the relative ratio of the gravitational binding energy to the total mass is in excess of $10^{-12} : 1$ (and not the miserable, unmeasurable ratio $10^{-47} : 1$ obtaining for the hydrogen atom).

To nobody's surprise – except perhaps to Dicke's – Einstein's strong equivalence principle proved to be correct. Dicke's own theory must be discarded, at least to all reasonable values of a new, adjustable parameter in his theory*.

To summarize, Kepler, Newton and Einstein each started with a different WHY for broadly the same set of phenomena. (To be more precise, Newton disclaimed any attempt at formulating a WHY for gravity theory - even though he apparently did build into it an equivalence hypothesis, justified later by Einstein's totally different approach.) Each theory gave predictions commensurate and better than the accuracies of the experiments then possible. However, at present, Einstein's approach remains the deepest - and the most accurately predictive - that we know of for explaining the existence - the raison behind - one of Nature's fundamental forces (gravity). Will this for ever be the case? Will this theory need modifications, extensions, become part of a bigger whole; will it even have to be discarded altogether, together with all its axiomatic sub-structure?

Einstein believed that the discovery of the deep WHY, underlying the other forces of nature will also follow the pattern of "geometrisation" of gravity that he had given to physics. Before I consider this, let me take one more example of differing styles of the offered WHYS at different epochs of physics. The example is from one of the other fundamental forces of Nature - electromagnetism. Maxwell, you may recall, predicted the existence of the electromagnetic radiation on the basis of the "displacement current" which he invented. This is one of the greatest feats of inventive discovery man has ever made - a discovery with few parallels, in the change it brought about in the world we live in. To-day an A-level student would demonstrate for you the necessity of a "displacement current" from the conservation law of electric charge. But Maxwell, himself, went through a tortuous - and what to-day might be considered an untenable - deduction based on a mechanical model of the aether. In Einstein's phrase, "(This) great

---

* Notice, like old soldiers, theories never die; they simply fade away. Thus, one could still save Brans-Dicke's theory, but only by assuming an outrageous value for this adjustable parameter. Other phenomena would then be affected but they are (hitherto) untestable.

change (was) brought about by Faraday, Maxwell and Hertz - as a matter of fact, half-consciously, and against their will - (because) all three of them, throughout their lives considered themselves adherents of the mechanical theory (of the aether).". Notwithstanding this, does anyone here tonight dare feel superior to Maxwell? Even after what I just quoted from Einstein, listen to his reverence for Maxwell: "Imagine Maxwell's feelings when the differential equations he had formulated proved to him that the electromagnetic fields spread in the form of polarized waves and with the speed of light. To few men in the world has such an experience been vouchsafed.".

§4. Consider now the forces of electromagnetism, and the two nuclear forces, weak and strong, responsible for radioactivity phenomena and for fission and fusion respectively. Recently, theory suggested and experiment confirmed that the weak nuclear force combines with electromagnetism - just as magnetism combined with electricity in the hands of Faraday and Maxwell a century ago - into one single, all embracing ELECTROWEAK force. The secret of this unification* lay in the extension of the so-called gauge ideas (well-known in electromagnetism) to the weak nuclear force. The characteristic of a gauge force is that such forces are proportional to the "charges" carried by the particles (e.g. $F = \frac{e_1 e_2}{r^2}$ for electromagnetism, $F = \frac{m_1 m_2}{r^2}$ for gravity).

---

* A crucial role in the demonstration of this electroweak unification was played by the ideas of "spontaneous" symmetry breaking. To motivate these, one has to invoke self-consistency (my second proposition, see §1) and to build in a special type of symmetrical potential into the structure of the theory - a potential which (surprisingly enough) yields solutions with less symmetry than what we started from. This potential should guarantee that the weak nuclear force remains short-range as observed, without affecting the long-range character of the electromagnetic force. There is a welcome price which one pays for inventing such a potential; one predicts the existence of a hitherto undetected particle - the so-called Higgs particle - which is currently being search for. This particle is welcome, for its existence would show that we are on the right track.

It is this sort of quantitative prediction, which distinguishes our use and our version of the self-consistency principle in physics, from empty philosophising.

What has been shown is that analogous to the electric charge, there exist three weak charges which determine the strength of the weak nuclear force and that these three charges - together with the electric - form four components of a "single" entity, each component transformable one into the other, through the operations of the group structure SU(2) x U(1) acting on an "internal symmetry space". I shall attempt to explain what I mean, more humanly, in a moment. But to complete the story: The future theoretical expectation is that the strong nuclear force is also a gauge force and the corresponding strong nuclear charges will eventually unite with the electroweak charges to make up a single entity, belonging to a still larger "internal symmetry group", of which the electroweak SU(2) x U(1) is a part*. From the concept of the electroweak force we shall, we hope, progress soon to the concept of a unified ELECTRO-NUCLEAR force, comprising electromagnetism as well as the two types of the nuclear force.

I have used the word "internal symmetry space" to designate that mysterious something which provides the present WHY for these unified gauge theories. Charge - electric, weak-nuclear, strong-nuclear - is a manifestation of the existence of an "internal" symmetry structure and of the postulated symmetries of laws of physics for rotations and other transformation in this mysterious internal space. The analogy of the internal space is with the familiar space-time. And the analogy of the electric and nuclear charges is with the gravitational charge - the inertial mass - whcih is associatзd with the translation-symmetry of the

---

Experiments to demonstrate this have just gone underway with Brookhaven-Irvine-Wisconsin and Milan-Turin-CERN-University College-Oxford collaborations. These are experiments designed to demonstrate that the proton is unstable with a half-life of the order of 10^{30} years. Hitherto the proton has been believed to be stable. (Compare 10^{30} years with the unmentionably tiny life of the Universe (of the order of 10^{10} years).)

four-dimensional space-time* continuum.

The question which arose in the nineteen-thirties when "internal symmetry spaces" were first invented by Heisenberg and Kemmer and which has become more and more insistent with the success of gauge ideas is this: Are these "internal spaces" purely mathematical constructs, or do they represent realistic adjuncts to the four dimensional space-time we are familiar with.

To take one example, one of the attempts currently being made is to describe physics in an 11-dimensional space-time. Of the 11-dimensions, four are the familiar space-time dimensions whose curvature is related to gravity and the other seven dimensions correspond to an internal symmetry space. In the theory advanced, the seven dimensions curled in upon themselves 10^{-43} secs. after the Big Bang, attaining a size of the order of 10^{-33} cms. and no more. We live on a cylinder in 11-dimensional space, our major source of sensory apprehension of these extra dimensions being the existence of charges - electric, weak-nuclear and strong-nuclear and the corresponding forces as manifestations of their curvature. Thus will Einstein's final dream (with which he lived for thirty-five years of uniting gravity) with the other (electronuclear) forces be eventually realized.

Exciting idea, which may or may not work quantitatively! But one question already arises; why the difference between the four familiar space-time dimensions and the seven internal ones? Why may the one lot curl in upon themselves, while the other does not? For the present, we shall make this plausible through the self-consistency principle; we shall invent a potential which will guarantee this as the only stable self-consistent dynamical system which can exist. There will be subtle physical consequences of this perhaps, in the form of remnants, like the black body radiation which was a remnant of the Big Bang. We shall

---

* Translation-symmetry is the statement that the laws of physics are independent of the location of where an experiment to test them is performed. This is one example of symmetry which we choose to ascribe to space-time structure; cf. the first proposition of §1. The experimental consequence of this assumed symmetry is the empirically testable conservation of energy and momentum.

search for these. Even if we find them, the next generation may perhaps question this entire mode of thought – particularly if a small discrepancy with our predictions is detected – and the cycle of questioning and answering might start all over again. Even to-day, an obvious question would be: Why eleven dimensions; why not a wholesome number, like thirteen? Or is this once again, due to the operation of the Bootstrap, the self-consistency principle?

There is an alternative suggestion to these extra dimensions which seeks to explain charges (other than gravitational) within the context of no more than a conventional four dimensional space-time. This suggestion, due to Wheeler, Schemberg and Hawking, does not add in new dimensions; it instead associates the electric and the nuclear charges to space-time topology – space-time Gruyere-cheesiness, worm-holes of the granular size of the order of 10^{-33} cms. The idea is attractive. Topology, you may recall, is concerned with "global" aspects as contrasted with the "differential" aspects of the present tradition in physics. It thus represents a real break with the past. Unfortunately – and I say this deliberately and ungratefully, in order to provoke some of my friends, in this audience – my own feeling is that the mathematics of topology, in respect of what we need, has not progressed beyond the Möbius strip and the Klein-bottle. Topology – as a language for physics is not yet capable of supporting the edifice the physicist may wish to erect on it. Could it be that our generation is defeated by the lack of development of a necessary mathematical discipline in a direction that we need? This has never happened before in the history of physics, but on this note, I would like to leave you to ponder on the deeper WHYS, appropriate to the physics of to-day – and tomorrow.