TEKST NR 27 1979

Helge Kragh

METHODOLOGY AND PHILOSOPHY OF SCIENCE IN PAUL DIRAC's PHYSICS

TEKSTER fra

IMFUFA ROSKILDE UNIVERSITETSCENTER

INSTITUT FOR STUDIET AF MATEMATIK OG FYSIK SAMT DERES FUNKTIONER I UNDERVISNING, FORSKNING OG ANVENDELSER

TEKSTER fra IMFUFA   Roskilde universitetscenter.

1/78 "TANKER OM EN PRAKSIS" - et matematikprojekt
Anne Jensen, Marianne Kesselhahn, Lena Lindenskov og Nicolai Lomholt.
Vejleder: Anders Madsen.

2/78 "OPTIMERING" - Menneskets forøgede beherskelsesmuligheder af natur og samfund.
Projektrapport af Tom J. Andersen, Tommy R. Andersen, Gert Kreinøe og Peter H. Lassen. Vejleder: Bernhelm Booss

3/78 "Opgavesamling", breddekursus i fysik.
Lasse Rasmussen, Aage Bonde Kræmmer, Jens Højgaard Jensen.

4/78 "Tre essays" - om matematikundervisning, matematiklæreruddannelsen og videnskabsrindalismen.
Mogens Niss.

5/78 "BIBLIOGRAFISK VEJLEDNING til studiet af DEN MODERNE FYSIKS HISTORIE"
Helge Kragh.

6/78 "Nogle artikler og debatindlæg om - læreruddannelse og undervisning i fysik, og - de naturvidenskabelige fags situation efter studenteroprøret"
Karin Beyer, Jens Højgaard Jensen, Bent C. Jørgensen.

7/78 "Matematikkens forhold til samfundsøkonomien"
B.V. Gnedenko.

8/78 "DYNAMIK OG DIAGRAMMER". Introduktion til energy-bond-graph formalismen.
Peder Voetmann Christiansen.

9/78 "OM PRAKSIS' INDFLYDELSE PÅ MATEMATIKKENS UDVIKLING"
Motiver til Kepler's: "Nova Stereometria Doliorum Vinarioum"
Projektrapport af Lasse Rasmussen.
Vejleder: Anders Madsen.

10/79 "TERMODYNAMIK I GYMNASIET"
Projektrapport af Jan Christensen og Jeanne Mortensen
Vejledere: Karin Beyer og Peder Voetmann Christiansen.

11/79 "STATISTISKE MATERIALER"
red. Jørgen Larsen.

12/79 "Lineære differentialligninger og differentialligningssystemer"
Mogens Brun Heefelt.

13/79 "CAVENDISH'S FORSØG I GYMNASIET". Projektrapport af Gert Kreinøe.
Vejleder: Albert Chr. Paulsen.

14/79 "Books about Mathematics: History, Philosophy, Education, Models, System Theory, and Works of Reference etc. A Bibliography".
Else Høyrup.

15/79 "STRUKTUREL STABILITET OG KATASTROFER i systemer i og udenfor termodynamisk ligevægt". Specialeopgave af Leif S. Striegler.
Vejleder: Peder Voetmann Christiansen.

HELGE KRAGH

METHODOLOGY AND PHILOSOPHY OF SCIENCE

IN

PAUL DIRAC'S PHYSICS

-oOo-

CONTENT

§ 1. INTRODUCTION.

The following essay is not intended to be a contribution to proper philosophy of science but rather a piece of history of science in a somewhat different key than what is usually seen. This key consists of a not too homogeneous mixture of aspects belonging to the history, philosophy and theory of science, all of them relating in some way or another to Paul Dirac's theories in physics.

A critical historical occupation with Dirac's physics needs no justification. Being one of the greatest Physicists ever, it is indeed strange that so little interest has been devoted to this scientist from the side of history of science. In particular, he is about the only one of the great generation of quantum pioneers whose work has never been examined from a methological and philosophical point of view. This task will fill out the main content of what follows.

To examine the entire scientific work of Dirac in a historical and philosophical context would be a very extensive task, and a pretty difficult one too. To facilitate things, I have chosen to concentrate on three areas of research which have a central placing in Dirac's conception of physics. These areas are the relativistic quantum theory, the theory of magnetic monopoles and the theory of cosmological constants. Together, these theories characterize fairly well Dirac's approach to physics and its involved philosophy. I do not think that a fuller inclusion of Dirac's theories, such as his important pre-1928 contributions to quantum mechanics, would materially change the picture here given.

The methological and philosophical elements in Dirac's physics will be examined largely in agreement with Einstein's famous advice: "If you want to find out anything from the theoretical physicists about the methods they use, I advice you to stick closely to one principle: don't listen to their words, fix your attention to their deeds" [67]. In doing so, it turns out that there is hardly one difinite method which can be called Dirac's. But it is possible to extract some invariant methological and philosophical themes from Dirac's scientific works, themes which more or less explicitly were played throughout his entire scientific career. Among these themes, the principle of plenitude and the principle of mathematical beauty are of particular interest and will be discussed in some details. In this connection, I shall take a less historical, and more general, view on aspects which are only indirectly linked to Dirac's physics. A special attention is directed to the role of mathematical beauty and other aesthetic factors in the process of obtaining scientific knowledge about nature, a question in which Dirac's position is shared by other prominent scientists. Philosophically, the discussion of this question is the core of the present essay.

For the benefit of mathematically trained readers, I have added three appendices where essential parts of the physical theories discussed in the text are outlined. The formulae herein may clarify some of the points made in the text but they are not necessary for an understanding of the main text.

§ 2. PAUL ADRIEN MAURICE DIRAC[1]

Though not well known to the public, Dirac is undoubtedly one of the greatest scientists ever, to be ranked on par with such giants as Newton, Maxwell, Einstein and Bohr. As most other pioneering scientists, Dirac created his fundamental theories at a young age. In 1933, aged 31 years, he received the Nobel Prize in physics.

Paul Dirac was born August 1902 in Bristol, England. At the age of sixteen he entered the University of Bristol, studying engineering sciences, and in 1921 he graduated in electrical engineering. In 1923 he went to Cambridge as a research student in 'applied mathematics', the then British synonym for theoretical physics, and there he encountered his first meeting with front research in physics. His research supervisor was Ralph Fowler, an expert on quantum theory and statistical mechanics. It was only late in 1923 that Dirac, from Fowler's lectures, learned about quanta and Bohr's model for the atom. Two years later, he had invented quantum mechanics! As a student in Cambridge, Dirac met with such top scientists as Larmor, Rutherford, Milne, Fowler and Eddington. Though Dirac was probably influenced by the thinking of these, and other Cantabrians, he had very little social contact with them. He preferred to spend his free time with long, solitary walks in the nature around Cambridge. Dirac stayed at Cambridge during his entire scientific career. Already in 1927 he became a Fellow of St. John's College, and in 1932, only 29 years old, he was appointed the Lucasian Professorship in mathematics, once held by Newton.

In an amazingly short time Dirac acquired a deep knowledge and a technical mastery of the front disciplines of current theoretical physics. Within a years stay in Cambridge, he had published his first scientific paper [23]; it dealt with the relativity dynamics of a particle and was submitted for publication through Eddington who took a keen concern in the gifted student. The break-through for Dirac came in 1925 when he successfully developed Heisenberg's fundamental ideas into a proper quantum mechanics, an abstract algebraic theory sometimes known as q-number algebra. This was done independently of the work taken place in Göttingen (Heisenberg, Jordan, Born) and, of course, also of Schrödinger's wave mechanics. In the following, we shall not be concerned with the first phase of Dirac's scientific life, i.e. the pioneering contributions from 1925-27. For these, we refer to Mehra's excellent survey.

Dirac's scientific career involves a variety of fields, all in theoretical physics. The following list includes the most important of Dirac's contributions:

Basic quantum mechanics. Development of quantum mechanical theory (1925-26). Various applications of quantum mechanics (1926-27). Quantum radiation theory (1927). Interpretation of quantum mechanics, transformation theory (1926-27). Quantum statistics (1926,1929).

Relativity quantum mechanics. Relativistic quantum theory for electrons (1928). Theory of electrons and positrons (1930-31). Generalized wave equation (1936). Positive-energy wave equation (1971).

Theory of magnetic monopoles (1931, 1948).

Cosmology. Theory of cosmological constants (1937, 1973).

Other subjects. Classical theory of the electron (1938, 1948). Quantum field theory. Hamiltonian dynamics. Gravitation theory. Theory of electrodynamics (1951).

Most of the great physicists contributing to the development of modern physics have had in common that their intellectual lifes were rich and versatile. Planck, Einstein, Bohr, Schrödinger, de Broglie and Pauli were all, in their own ways, deeply interested in matters outside physics, whether being philosophy, society, art, politics or literature. Also, most of the pioneers were socially minded personalities. Some of them, like Sommerfeld, Bohr and de Broglie, gave rise to schools in science.

Dirac was an entirely different kind of personality. He always was an introvert person, favouring isolation rather than social or intellectual contacts. Of his childhood, Dirac's biographer, Jagdish Mehra, reports that, "The need for social contacts was not much emphasized in his home. Paul was an introvert and, being often silent and alone, he devoted himself to the quiet contemplation of nature" ([122], p.18). A loner, a prototype of the individual genious, Dirac has never cared about cooperation in physics. Out of Dirac's entire stock of publications, covering some 130 titles, only five were written jointly with other authors. The highly original tenor in Dirac's scientific works is also reflected in his small use of references. In some contrast to what is normal for scientific publications, Dirac cited only few authors, as he really had no need to do so; the average number of cited works in Dirac's most important papers was about five or six.

During his age of growth, Dirac did not meet with a particularly 'intellectually stimulating climate', nor did he take part in social or philosophical debates of any kind. In contrast to bourgeois education on the Continent, where languages, literature and classical subjects were highly ranked, Dirac's early education did not include these many-sided aspects of learning but concentrated on physics, mathematics and chemistry. In school Dirac never read Plato, and he has hardly done so since then.

To a remarkable extent, Dirac concentrated his talents one-sidedly on physics. Schrödinger, who was a loner like Dirac but with a very different background and with wide intellectual interests, once wrote: "Physics consists not merely of atomic research, science not merely of physics, and life not merely of science."[2] For Dirac, however, life was physics. Dirac's contemporary, the German physicist Walther Elsasser, has characterized his British colleague in the following way:

> "He had succeded in throwing everything he had into one dominant interest. He was a man, then, of towering magnitude in one field, but with little interest and competence left for other human activities. ... In other words, he was the prototype of the superior mathematical mind; but while in others this had coexisted with a multitude of interests, in Dirac's case everything went into the performance of his great historical mission, the establisment of the new science, quantum mechanics, to which he probably contributed as much as any other man." ([70],p.51)

In the twenties and thirties, most influential physicists in England as well as on the Continent were involved in discussing also religious views and their relationship to the new physical world-picture. Most of the leading physicists were affected by religion, although rarely by orthodox christianity. Eddington , for instance, was deeply influenced by his quaker background and often attempted to merge his philosophy of physics with his religious views. Dirac seems to have been quite outside this religious trend, rather favouring an atheist or maybe agnostic view. Heisenberg ([93];pp.91-93) recalled discussions on religion during the Solvay Congress in 1927 where Dirac definitely rejected any religious idea. Religion, Dirac said at this occasion, is a system of myths with an ideological function, an opium for the people. Scientifically viewed, it is based on irrationality and false postulates and hence without any appeal to the scientific man, Dirac argued. Wolfgang Pauli commented on Dirac's view with usual sarcasm: "But yes, our friend Dirac has a religion, and the basic postulates of this religion is: 'There is no God, and Dirac is his prophet'".

Dirac's concern with art and literature was always modest, if not hostile. "How can you do physics and poetry at the same time?" he once asked Oppenheimer, "the aim of science is to make difficult things understandable in a simple way; the aim of poetry is to state simple things in an incomprehensible way. The two are incompatible."[3] With the exception of cosmology, Dirac has never attempted to change his interests towards other sciences, or to apply his physical insight to e.g. biology, such as became a fashion at many a quantum pioneer.

Dirac's attitude to poetry, religion and non-scientific matters in general, may call in mind earlier periods of rationalism when hostile attitudes to poetry were held by e.g. Newton, Locke and Hume. In many ways, indeed, Dirac's general outlook seems strikingly similar to the outlook which dominated the Newtonian Age of Reason. Dirac might have felt himself more at home with the spirit of this age than with his own.

Also regards philosophy, Dirac's explicit concern was virtually nil. While in Bristol, Dirac felt at a time that philosophy might perhaps be important and he tried to do some reading in philosophy. He followed C. D. Broad's lectures on the philosophical aspects of relativity [10] and read Mill's Logic [128] and a few other books. However, Dirac's meeting with the philosophy was short and unsuccessful. "I felt then that all the things that philosophers said were rather indefinite, and came to the conclusion eventually that I did not think philosophy could contribute anything to the advance of physics"([53],p.111). This attitude seems to have remained at Dirac throughout his entire scientific life.

In England the twenties and thirties were years with a widespread interest in natural philosophy and in the philosophical implications of the new physics. A stream of books and articles were published on aspects of modern physics in an epistomological context, the most important due to Russell, Joad, Whitehead, Eddington and Broad. Dirac's reasoning in physics may have some affinity with current philosophical themes in England in these decennia. But it is doubtful whether Dirac was influenced to any extent by these philosophers, with the possible exception of Arthur Eddington.

Ever since his days as a student in Bristol, Dirac was fascinated by Eddington's approach to physical science, based on a peculiar mixture of grand syntheses, epistomological principles and advanced mathematical techniques. In particular, Eddington was the British fountainhead in relativity, a field which always had a special appeal to Dirac. At an early age, Dirac was caught up in the general excitement over relativity which swept over Europe in the post-war years. While at Bristol, he read much on relativity theory, in particular Eddington's <u>Space, Time and Gravitation</u> which made a great impression on the young student. Ever since, Dirac has been firmly commited to relativity which he always regarded as the ideal of a physical theory. Dirac has only published little on relativity theory but a large part of his work is, directly or indirectly, connected with Einstein's theory. His devotion to relativity was, as we shall see, important to his general views concerning the aims and methods of physical science.

The professional relationship between Eddington and Dirac was one of mutual impact. Eddington was much impressed by his young colleague's approach to quantum mechanics. The relativistic electron theory of 1928 made a particularly strong impact on Eddington who at many occasions attempted to generalize Dirac's theory and to connect it with cosmology. This ambitious programme culminated in 1936 with what probably was Eddington's opus magnum, the monumental but ill-fated Theory of Electrons and Protons. Also from an earlier date Eddington had thought Dirac's physics to be congruent with his own philosophical views in which the physical universe was pictured as a dematerialized world of shadows dressed up in mathematical symbols. In his 1928 classic, The Nature of the Physical World, based on lectures delivered in January to March 1927, Eddington praised Dirac's q-number algebra for its consequent symbolism and emancipation from visualizable models:

> "If we are to discern controlling laws of Nature not dictated by the mind it would seem necessary to escape as far as possible from the cut-and-dried framework into which the mind is so ready to force everything that it experiences. I think that in principle Dirac's method asserts this kind of emancipation."([64],p.210)

Eddington's concern with quantum mechanics was, however, unorthodox and his ideas on the subject were in gross dissonance with what most quantum physicists, including Dirac, regarded as proper quantum mechanics ([172],p.83ff;[124],p.112ff). At one occasion Dirac felt obliged to protest publicly against Eddingtons's critique of current standards in quantum mechanics. Referring to Eddington's objection aginst the customary use of Lorentz transformations in quantum mechanics, Dirac mildly corrected his senior colleague: "The issue is a little confused because Eddington's system of mechanics is in many important respects completely different from quantum mechanics, and although Eddington's objection is to an alleged illogical practice in quantum mechanics he occasionally makes use of concepts which have no place there" ([39],p.193).

As one of the founders of quantum mechanics and a leading contributor to its physical interpretation, one might expect that Dirac was also engaged in the debate over the philosophical implications of quantum mechanics, such as were most other pioneers. But Dirac has never contributed to quantum philosophy, so-called. Dirac's stand in the debate over objectivity in physics and kindred questions, eagerly discussed in the thirties, was one of superior indifference. Some physicists, like Jordan [106], rejected the problem of an objectively existing world as being meaningless, in the positivistic sense. Dirac was not a positivist, he just didn't care about ontological problems. Concerning the quantum mechanical measurement process and the existence of an objective nature, Born portrayed in 1936 what he called Dirac's l'art pour l'art attitude:

> "Some theoretical physicists, among them Dirac, give a short and simple answer to this question. They say: the existence of a mathematically consistent theory is all we want. It represents everything that can be said about the empirical world; we can predict with its help unobserved phenomena, and that is all we wish. What you mean by an objective world we don't know and don't care."([7],p.12)

Such a standpoint is, as Born remarked, scientifically unobjectionable but it is restricted to a small circle of experts. While this was a touchstone to Born and others who wanted to make science intelligible to every thinking man, the aristocratic Dirac didn't share such worryings. Science for the people has never been on Dirac's programme.

In the post-1927 division of scientists and philosophers in the Copenhagen school (Bohr, Heisenberg a.o.) and the 'deterministic' school (Einstein, Schrödinger a.o.), Dirac rather belonged to the former, though not caring much about the debate. He has always believed that the Copenhagen interpretation is acceptable as the best one in the state of present knowledge; but also that our present foundation of quantum mechanics is not the last word and probably will have to be replaced by some fundamentally new theory in the future.

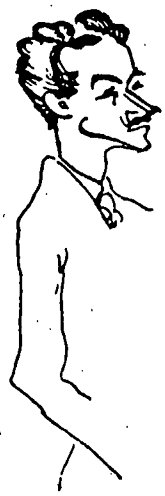
This brief sketch shows that whatever 'philosophy' may be connected with Dirac's physics, it can hardly be attributed to earlier philosophical influences, intellectual longings or cultural contacts. Among the great physicists, Dirac was probably the least philosophically minded ever. This does not mean, of course, that Dirac's physics is devoid of philosophical elements or that 'philosophy' did not enter his scientific reasoning. It certainly did, as well as we shall see. But it means that Dirac's philosophy of science must be examined in a different way than e.g. the philosophy of science of Bohr, Schrödinger or Weyl. To a unique extent, the methodological and philosophical content of Dirac's physics is unconnected to external factors. It is caused in the physics itself; or rather in Dirac's unsophisticated and unphilosophical reflections over his physical theories. Certainly no scientist can avoid somehow to be influenced by forces outside his scientific milieu. Even the greatest scientists are, as Schrödinger emphasized, members of a social and cultural environment:

> "The scientist cannot shuffle off his mundane coil when he enters his laboratory or ascends the rostrum in his lecture hall. In the morning his leading interest in class or in the laboratory may be his research; but what was he doing the afternoon and evening before? He attends public meetings just as others do or he reads about them in the press. He cannot and does not wish to escape discussion of the mass of ideas that are constantly thronging in the foreground of public interest, especially in our day. Some scientists are lovers of music, some read novels and poetry, some frequent theaters. Some will be interested in painting and sculpture .... In short, we are all members of our cultural environment."([154],p.98)

Schrödinger's observation of the influence of the cultural environment on scientists' thinking is, no doubt, correct. Although it is not restricted to the fine arts but also covers economical, political and ideological influences, not considered by Schrödinger. In Dirac's case, however, this influence was weak. This implies, futhermore, that it seems rather hopeless to apply a socio-cultural scheme of explanation, such as suggested by Forman [79] and Feuer [74] in different versions, to Dirac's case. Feuer has tried to do so (see §7) but, as one would expect, with little luck. That externalism, in whatever form it may take, seems unable to cope with Dirac's physics, is not, I hardly need to say, an argument against externalistic historiography as such.

PAUL DIRAC IN GEORGE GAMOW's LINE AND VERSE

(From Gamow's Faust parody, written to Niels Bohr's Institute for Theoretical Physics in 1932. See [82])

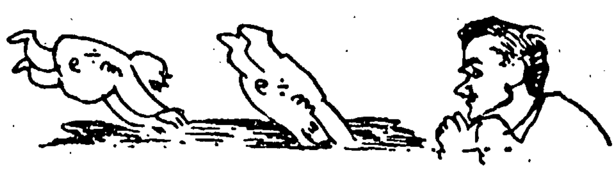


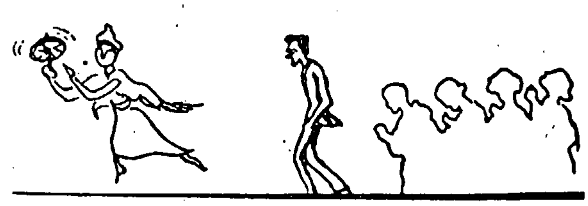
(DER MONO-POL *tritt auf, singt*):

Es waren zwei Mono-Pole,[30]
Die hatten cinander so lieb.
Sie konnten zusammen nicht kommen,
Denn Dirac war allzu tief.

DIRAC:

That donkey-electrons should wander
Quite aimless through space, is a slander,
That only with articles
On hole-like particles
Could be said to have found a defender.





§ 3. A DIGRESSION ON AESTHETICAL PRINCIPLES IN THE HISTORY OF IDEAS.

In the course of history, intellectual work of any sort has been intimately linked to various considerations of what may be called, with a not too fortunate name, aesthetical principles. This is manifest in the history of art and literature but also in the history of science such considerations have played a central, though not always noticed, role. The interaction between aesthetic and scientific concepts is, futhermore, not merely a feature of past science. Although usually disguised, it is present also in modern science, as exemplified in Dirac's physics.

The role of aesthetical or trans-logical principles in modern science has recently been studied by Lewis Feuer [75]. He regards the use of what he terms 'teleological principles' as a kind of objectivization of some underlying emotional aim: "A particular kind of world is sought which will answer to the scientist's emotional longings. A teleological principle, in its most general sense, is one which affirms that some ethical, extra-logical purpose is fulfilled in the structure of the laws of nature. Such a principle, moreover, serves then as a heuristic agent for discovering those laws of nature. It is not an after-the-fact theological commentary but an active participant in the work of exploration" (ibid., p.378). The sources for fundamental scientific inquiry have then to be traced back to the individual scientist's psychology, to his emotional longings. These longings are expressed in a teleological principle which is 'isomorphic' with the scientist's ethical, social or religious world-view, his emotional <u>a priori</u>. Feuer is, I think, quite right in his appraisal of such principles as being essential to the functioning of science on its fundamental level. But apart from the dubitable association of teleological principles with emotional aims - and this is the whole core of Feuer's thesis - 'teleological' seems to me an unfortunate name. In some cases, principles of this kind work indeed actively as heuristic agents, and may then be described as teleological. In other cases, however, the same principles are really "after-the-fact theological commentaries", i.e. rationalizations a posteriori and not a priori research programmes. This holds in particular for the principle of beauty in science, such as we shall see in the latter sections of the present study.

In the following, I shal use the term aesthetical principles in a somewhat less comprehensive way than Feuer's teleological principles. Still the class of aesthetical principles is extremely varified. It covers such items as plenitude, beauty, simplicity, continuity, harmony, symmetry and invariance. A fuller treatment of the role of these principles in the history of science is out of the scope of the present work. For our purpose a brief introduction of selected aspects, to appear in Dirac, will suffice. A further discussion will be taken up in §9 and §10.

First some comments on the principle of plenitude, largely based on Lovejoy's treatise [117]. Cooked down to its essence, the principle of plenitude states that 'every ontological niche must be filled', or, less obscurely, that what is conceived to be possible is also somehow to be endowed with physical reality. During the course of time this theme has been exposed in many different versions, with different purposes and with different effects. Appearing first in Plato, the principle of plenitude has played a central role in the formation of ideas in theology, ethics, philosophy and science. Particularly in the 18th century it was adopted and developed by Spinoza and Leibniz as well as by many authors of a less imposing stature. To these philosophers it was tightly linked to questions of theology, especially to the question of God's omnipotence. Spinoza's version of the principle of plenitude was that "whatever we conceive to be in the power of God necessarily exists" (ibid.,p.152). To Leibniz the principle was closely related to the principle of continuity, that 'nature makes no leaps,' and that, therefore, the order of natural quantities forms a single chain in which the various elements exhaust the space of potential being. This belief in fullness and continuity sometimes led to 'predictions' and entire research programmes in the sciences, particularly in anthropology and biology. Leibniz thus argued that intermidiate links between animals and plants, zoophytes, have to exist in nature.

> "So great is the force of the principle of continuity, to my thinking, that not only should I not be surprised to hear that such beings had been discovered ---- but, in fact, I am convinced that there must be such creatures, and that natural history will perhaps some day become acquainted with them." (ibid.,p.145)

From Leibniz' occupation with the principle of plenitude, Feuer has proposed to call it 'Leibniz' principle'. This, however, is not a quite fortunate choice since Leibniz was notoriously vague on the point and did not take the principle as an absolute. He held, for instance, that there are species which are possible but nevertheless do not exist.

A more precise definition of the principle of plenitude was given by a later 18th century author, the French philosopher J. B. Robinet, who put it in these words: "From the fact that a thing can exist I infer readily enough that it does exist" (ibid.,p.272). Now, if the principle of plenitude shall be a useful scientific instrument, the problem is, of course, to define what is meant by the ambiguous phrase 'a thing <u>can</u> exist'. Surely, this potential existence should not be identified with imagined existence. Centaurs are imaginable in fantasy, byt their existence is not in the physical universe. The trouble with a useful formulation of the principle of plenitude was accentuated with the romantic wave, where even errors, disharmony and irregularity were included in the necessary richness of the world. Exposing the romantic version of the principle of plenitude, Schiller wrote:

> "Every offspring of the brain, everything that wit can fashion, has an unchallenged right to citizenship in this larger understanding of the creation. --- That great Householder of his world ... could not permit even error to remain unutilized in his great design, could not allow this wide region of thought to lie empty and joyless in the mind of man ..." (ibid.,p.299)

Of the more scientific applications of the principle of plenitude, the periodic system furnishes an early example. Mendeleev and others predicted successfully that new elements would exist from the sole argument that this would be in harmony with the required periodicity of the elements. And later Rydberg, less successfully, predicted other elements to fill out the 'ontological niches' which he thought prescribed by the periodic law. In theoretical physicists' use of the principle of plenitude, the ambiguous phrase 'can exist' in Robinet's formulation is, it turns out, rather identified with 'consistencey with the fundamental laws of nature.' In Dirac, this principle found a virtuoso.

Better, perhaps, we may express the functioning of the principle of plenitude in modern physics by saying that entities are asumed to be realized in nature as far as they are described mathematically consistent and are not ruled out by a so-called principle of impotence. These principles are general statements, asserting the impossibility of achieving something ([169],p.58ff). They are exemplified in the second law of thermodynamics, the uncertainty principle in quantum mechanics and in the impossibility of recognizing absolute velocity in the theory of relativity. It is a widespread belief that ultimately the basic laws of physics may all be expressed as similar postulates of impotence. Such postulates are not forced upon by logical necessity; universes in which any postulate of impotence is violated are intelligible, and in fact readily so. In this respect, they have an empirical basis. But they are not, on the other hand, simply generalizations of experimental facts. They are the assertion of a conviction, guided by experience indeed but raised to an a priori status, that all possible attempts to do a certain thing are bound to fail. In this respect, postulates of impotence share the characteristics of other aesthetical principles in science.

Beauty is another of those aesthetical, a priori principles which so thoroughly have penetrated all aspects of the history of creative thinking. In a scientific context the quest for beauty was a predominant theme in e.g. Bruno and Kepler. The latter, for instance, at one occasion argued for the heliocentric system by this purely aesthetic confession: "I certainly know that I owe it this duty, that as I have attested it as true in my deepest soul, and as I contemplate its beauty with incredible and ravishing delight ..." ([110],p.116). In modern time the quest for beauty in science has been highlighted particularly by Poincaré and his school of philosophy. "The scientist", Poincaré said, "does not study nature because it is useful to do so. He studies it because he takes pleasure in it; and he takes pleasure in it because it is beautiful. ... It is because simplicity and vastness are both beautiful that we seek by preference simple facts and vast facts ..." ([141], p.22). Closely following Poincaré, Hadamard considered invention in mathematical sciences, including theoretical physics (in which case non-conventionalists would prefer to speak about 'discovery' instead of 'invention'), as essentially a choice among various combinations of thought. This choice is, and should be governed by affective elements, particularly by the scientists' sense of mathematical beauty. This feeling - call it beauty or not - is, in fact, the only guide for basic scientific investigation on which the theoretical researcher can possibly rely. Hadamard states: "The guide we must confide in is that sense of scientific beauty, that special esthetic sensibility ..." ([86],p.127).

'Beauty' may seem to be a somewhat misleading term in a scientific context, since the word is usualle associated with the experience of appearances in nature, in art and the like. Poincaré takes pains to explain that beauty in science is quite different from the sensational beauty connected with artistic experiences. Science, Poincaré emphasizes, is intellectual work, dealing with "that more intimate beauty which comes from the harmonious order of its parts, and which a pure intelligence can grasp" ([141],p.20).

Following Poincaré, the principle of scientific beauty has got its modern philosophical advocate in Michael Polanyi, to whom the sense of beauty in science is essential in the context of discovery, as a heuristic means, as well as in the context of justification. "A great scientific theory", Polanyi says, "has an inarticulate component acclaiming its beauty, and this is essential to the belief that the theory is true. No animal can appreciate the intellectual beauties of science" ([142],p.133). And, "We recognize intellectual beauty as a guide to discovery and as a mark of truth" (ibid.,p.300).

Even if the basic scientist, according to Poincaré and others, is preoccupied with aesthetic qualities rather than with utility and empirical truth, these are not foreign to the programme of beauty in science. On the contrary, the point is that exactly those theories which satisfy our sense of beauty are likely to be true and useful (although here utility is to be taken in a strict intellectual sense and has nothing to do with social or industrial utility). This identification of truth with beauty, central to the whole idea of beauty in science, was many years earlier expressed poetically by Keats:

> "Beauty is truth
> truth beauty - that is all
> Ye know on earth
> and all Ye need to know."[4]

One and a half centuries later, Dirac restated Keats' confession in another context: "A theory with mathematical beauty is more likely to be correct than an ugly one that fits some experimental data" ([49],p.29). To judge truth by criteria of beauty may then, in some quarters, be considered the task of science. The converse thesis, that the truth embodied in intellectual work is relevant to its aesthetic merit, plays a role in the philosophy of art [127] but need not concern us here.

Closely connected to the principles of beauty and plenitude are <u>the principle of simplicity</u> and what may be considered as its mature relative, <u>the principle of least action</u>. Indeed, simplicity is most often taken to be an important ingredient of scientific beauty. The principle of plenitude and the principles of simplicity and least action have common historical roots and to a large extent they have developed parallelly. Both were central tenets in Leibniz' philosophy where they were, furthermore, tied up with moral and theological issues. This alliance gradually weakened, but it never disappeared. Two hundred years after Leibniz, Planck undertook to argue for the existence of God from the principle of least action ([119],p.165).

Although belonging to the same class of aesthetical and teleological suppositions, minimum principles do not share the ambiguity and arbitrariness of the principles of beauty and plenitude. Since the days of Maupertuis and Fermat, the principle of least action has been developed into scientifically exact formulations, which have proven immensely effective in the advancement of science.

At some periods, certain aesthetical principles have been conceived not merely as useful heuristical guides in science, but as universal canons to dominate the entire knowledge of nature. Among the great physicists, Helmholtz and Planck both held that all of physics has to be subordinated principles of simplicity and least action, considered to be ends, not only means of scientific inquiry. To Helmholtz the principle of least action was "the universal law pertaining to all processes in nature" ([119],p.163). And Planck, in 1915, saw the modern revolution in physics as providing support for his belief in the universality of minimum principles:

> "The most brilliant achievement of the principle of least action is shown by the fact that Einstein's theory of special relativity, which has robbed so many theorems of their universality has not disproved it, but has shown that it occupies the highest position among physical laws. --- The principle of least action ... appears to govern all reversible processes in Nature." ([83], 140)

There are, of course, quite a number of other aesthetic principles of importance in science. Principles of symmetry and invariance, playing such a profound role in modern physics [170], are thus belonging to the same class as the principles mentioned above. Like simplicity, they are often considered to be constituents of beauty in science. And like the principle of least action, they have developed from qualitative, ambiguous principles to become quantitative and highly effective instruments of physics. Lastly, aesthetics in science is sometimes conceived more broadly, covering, for instance, individual scientists' views of *ad hoc*'ness, their ideas of visualization and imagery, their modes of thinking and choice between opposing themata of any kind [95],[126].

In discussing aesthetical principles in science, one should take care to distinguish between 'objective aesthetics' and 'subjective aesthetics'. In the first case it is held, for instance, that Nature posesses an immanent tendency to simplicity or beauty; and the pragmatic success of these principles simply to be due to the fact that the laws of Nature - or God's mind - *are* simple or beautiful. The second case is independent of any objective structure of the universe; it demands that in *interpreting* nature the scientist should prefer simple explanations to complex, beautiful equations to ugly, invariant laws to non-invariant etc. Of course, these two meanings are often mingled together, since the mathematical description of nature is often conceived to be a, more or less accurate, reflexion of how the nature is really constituted. This unification may be clearly seen in Newton's famous 'first rule of reasoning in philosophy': "We are to admit no more causes of natural things than such as are both true and sufficient to explain their appearence;" for, Newton says, "Nature does nothing in vain and more is in vain when less will serve; for Nature is pleased with simplicity, and affects not the pomp of superfluous causes" ([134],p.398).

In other philosophical conceptions, however, the unification is unwarranted. This holds foremost in phenomenalistic and conventionalistic philosophies of science, where physical knowledge consists solely in an economic systematisation of relations between phenomena, whose connection to an objective nature is declared meaningless to discuss. Mach's principle of thought economy is indeed a principle of simplicity, considered to be a prerequisite for any scientific knowledge, but it is totally disconnected from any ideas of simplicity inherent in nature. It has a psychological meaning, not an ontological. This point, contrasting Newton's, has been expressed by many scientists. Max Born, for instance, commented in 1939 on the possible unitary theory of the future in the following way:

> "... wir dürfen vermuten, dass, wenn wir sie finden, sie die Form eines Extremalprinzips haben wird, nicht weil die Natur selber einen besonderen Willen hat oder Zweck verfolgt oder besonders ökonomisch ist, sondern weil der Mechanismus unseres Denkens keinen anderen weg kennt, um der komplizierten Struktur der Naturgesetze kurzen, präzisen Ausdruck zu verleihen." ([9],83)

§ 4. RELATIVITY THEORY OF ELECTRONS AND POSITRONS.

Dirac's fame in physics rests, above all, on his work with the negative and positive electron, completed while he was still in his twenties. It was for this work that Dirac received the Nobel Prize in 1933.

I have subjected Dirac's relativistic quantum theory of electrons to a historical investigation elsewhere [113] and shall therefore suffice to outline it in brief. The problem in focus, was how to describe mathematically the quantum mechanical behaviour of an electron so as to be in accordance with the principle of relativity. Originally, quantum mechanics was presented in a non-relativistic way and consequently many physicists attempted to formulate relativistic generalizations of the new physics. These attempts took two directions, either by adding relativity as a correction to the non-relativistic theory or by straightforwardly extending the methods of quantum mechanics into the relativistic domain. In the first case, this led to a theory which accounted well for spin and relativity phenomena in a first approximation, which was sufficient to match the experimental accuracy; but the theory was not fully relativistic, i.e. not Lorentz invariant. The second case led to the so-called Klein-Gordon equation, which was tailored to be Lorentz invariant but failed to explain spectroscopic phenomena involving spin.

Confronted with the problem of reconciling quantum mechanics and relativity, Dirac approached the matter in a highly original and remarkable way, rejecting both of the mentioned alternatives. Dirac's solution to the problem was offered in the beginning of 1928 [25]. It was in many ways typical for Dirac's particular way of thinking about physical problems. What prompted Dirac to repudiate the Klein-Gordon approach as well as the semi-empirical approach followed by Darwin, Pauli and others, was his conviction that "the interpretation of the relativistic quantum theory [should be] just as general as that of the non-relativity theory" (ibid.,p.612).

This emphasis on generality was always a significant feature in Dirac's methodology. Almost forty years later, he described the aim of theoretical physics as follows:

> "Our object is to get a single comprehensive theory that will describe the whole of physics ... We need a single consistent comprehensive theory. Any special theory that one sets up for dealing with a particular problem should be consistent with this general theory." ([47],pp.1-2)

The emphasis on generality and comprehensiveness did not imply, however, that Dirac was methodologically devoted to strict deductivism or that he favoured grand syntheses to piecemeal improvements. "One should not try to accomplish too much in one stage," Dirac has said, "One should separate the difficulties in physics one from another as far as possible, and then dispose of them one by one" ([48], p.24).

In 1928, the principle of generality implied that the quantum theory, sought for, had to conform to the general requirements of relativistic and quantum mechanical transformation theory. This again means that the wave equation has to be Lorentz invariant as well as of the first order in the time derivative (see also Appendix I). These principles - principles of faith in the foundation of fundamental physical theory - were a couple of years later formulated in this way:

> "The important things in the world appear as the invariants (or more generally the nearly invariants, or quantities with simple transformation properties) of these transformations ... The growth of the use of transformation theory, as applied first to relativity and later to the quantum theory, is the essence of the new method in theoretical physics." ([29],p.V)

Although written two years after Dirac's success with the relativistic wave equation, it expresses well Dirac's motives also in the creation of this theory.

Guided by his belief in these principles, Dirac managed to construe a quantum mechanical theory which in a very satisfactory way accounted for all the relevant experimental experiences of the time. Dirac's way to the theory was markedly different from the attempts of other physicists. First of all, the approach and the motives were basically non-empirical: the experimental anomalies played but a remote role for Dirac who was primarily concerned with formulating a consistent theory from methods of principles. Therefore the approach taken by Darwin and others had no appeal to Dirac. This feature was recognized by Darwin who acknowledged that "Dirac's great success in finding the accurate equations shows the great superiority of principles over the previous empirical method" ([19],p.664). It would be but a minor exaggeration to claim that Dirac might have found his equation even if he had no knowledge at all of the current development taking place in empirical physics. In this respect, as in others, Dirac's approach was close to Einstein's (see §9). It may be difficult to characterize Dirac's 1928 approach in conventional methodological labels, but whatever it was it was not an example of the inductive-empirical method.

In his later years, Dirac has often stressed that in the theoretical sciences, mathematics is to be the leading part (see §8). In 1928, Dirac had not yet developed such as kind of philosophy. Yet, traces of it may be found implicitly in his relativistic theory. In the course of arguments leading to the Dirac equation, it thus turns out that the wave function under consideration must, as a consequence of mathematics, be one with four components, $\psi = (\psi_1, \psi_2, \psi_3, \psi_4)$. Dirac accepted this result although it was not at all supported by empirical evidences. To describe an electron, only a two-component wave function is required, such as the one introduced by Pauli in 1927 in order to account for the two directions of the spin. In Dirac's theory, then, half of the components were 'superfluous', without any physical justification whatsoever. This should be a disqualifying feature according to the standards of Ockham's Razor, usually regarded to be a sound methodological principle also in modern science. Still Dirac felt compelled to follow the logic of his mathematical reasoning, happily unconcerned about Ockham's rule.

Contrary to what is usually associated with 'the British spirit' in physical science, Dirac firmly rejected visualizable models. If used heuristically models are, according to Dirac, likely to lead to false results; and they are, furthermore, in principle opposed to the nature of physical reality.

> "The main object of physical science is not the provision of pictures, but is the formulation af laws governing the phenomena .... whether a picture exists or not is a matter of only secondary importance .... There is an entirely new idea involved, to which one must get accustomed and in terms of which one must proceed to build up an exact mathematical theory, without having any detailed classical picture." ([29],p.12)

In the preface to his classic textbook Principles of Quantum Mechanics, Dirac stated in an apparently Eddington inspired passage:

> "Nature's fundamental laws do not govern the world as it appears in our mental picture in any very direct way, but instead they control a substratum of which we cannot form a mental picture without introducing irrelevancies." ([29],p.V)

The sharp distinction between the familiar "world as it appears" and an underlying "substratum" was a typical theme in Eddington as well as in other contemporary scientist-philosophers. The Eddingtonian spirit in the 1930 preface is even clearer in the continuing statements. Here Dirac said that the current state of affairs in theoretical physics "is very satisfactory from a philosophical point of view, as implying an increasing recognition of the part played by the observer in himself introducing the regularities that appear in his observation" (ibid). This declaration is congenial to conventionalism in general and to Eddington's so-called selective subjectivism in particular. According to this idea, our observational knowledge is selected by our sensory and intellectual equipment and is therefore partly subjective; the phenomena of nature are forced to fit into the observer's sensory and instrumental settings, are partly manufactured instead of discovered (cf.[124] and [172]).

It is true, as pointed out by Mehra ([122],p.50), that 'pictorial models' are predominant in Dirac's thinking, often used in order to think about premature physical concepts and to transform vague ideas into mathematical symbols. However, the pictorial models used by Dirac have almost nothing in common with the traditional models in physics, functioning essentially on classical lines. "One may," Dirac stated, "extend the meaning of the word 'picture' to include any way of looking at the fundamental laws which makes their self-consistency obviuos. With this extension, one may gradually acquire a picture of atomic phenomena by becoming familiar with the laws of the quantum theory" ([29],p.12). This is, of course, an entirely different conception of models than the one adopted by the classical models makers of the Victorian Age, such as Lord Kelvin, J.J.Thomson and J.Larmor.

In accordance with his anti-model standpoint, and in contrast to most other physicists working in the field, Dirac consciously preferred to look upon the electron as a point charge so that questions concerning any model for its internal structure were ruled away. This standpoint is a sort of application of a simplicity principle. Commenting on the semi-classical spin models of the electron, picturing it as a rotating ball, Dirac asked: "The question remains as to why Nature should have chosen this particular model for the electron instead of being satisfied with the point-charge" ([25],p.610). A similar kind of simplicity argument was used by Dirac when he later engaged in the classical theory of electrons. Dirac again treated them as point charges, arguing that "it seems more reasonable to suppose that the electron is too simple a thing for the question of the laws governing its structure to arise" ([37],p.148). The use of 'Nature' as a heuristic principle, also explicitly applied by Dirac at other occasions (see §5), closely resembles Einstein's use of 'God'. Einstein used to ask himself *e.g.* whether God could have created the world in this or that way. Dirac, the atheist, addressed the same kind of questions to Nature.

When Dirac created his new theory of electrons, it was welcomed as a revolutionary step forward. Very soon, however, serious objections were raised against Dirac's theory. These objections were above all concerned with the problem of negative energy states.[5]

Already in 1928, Dirac had clearly recognized this problem although at that time leaving it untreated. The case is, briefly stated: the solution af the Dirac equation (and, in fact, of any relativistic theory) allow for negative total energies (i.e. $E< -m_oc^2$) in addition to the usual positive energies. That is, if solved completely the Dirac equation results in a class of solutions which formally may be attributed to 'electrons' with a positive charge and a negative energy. These solutions, it turns out, are associated with the 'superfluous' components of the wawe function, mentioned above (see appendix I). Dirac considered this peculiar situation in the words:

> "One gets over the difficulty on the classical theory by arbitrarily excluding those solutions that have a negative W [total energy]. One cannot do this on the quantum theory, since in general a perturbation will cause transitions from states with W positive to states with W negative. Such a transition would appear experimentally as the electron suddenly changing its charge from -e to e, a phenomenon which has not been observed." ([25], p.612)

Now, particles with negative energy are quite unintelligible in physics,[6] as they lead to absurd consequences; they will for instance, have less energy the faster they move and will have to absorb energy in order to be brought to rest.[7] In 1928, Dirac chose to ignore the negative energy states on the ground that they would give rise to phenomena "which have not been observed" ([25],p.612). Still he realized that if the formal power of the theory should be preserved, the negative energy solutions had somehow to be taken seriously. Without an understanding of the negative energies, the theory was doomed to remain "still only an approximation" (ibid). And Dirac's aim was an exact theory.

During the following years, the negative energies caused much concern among the quantum physicists. In the beginning of 1929, Heisenberg wrote about "die Elektronentragödie nach Dirac (+mc → -mc)", and Pauli was equally disturbed over the "schauerlichen Konsequenzen der beiden Vorzeichen für die Elektronenmasse."[8] Some physicists, and Schrödinger in particular, tried to reconstruct Dirac's theory so as to get rid of the negative energy solutions. Dirac, however, thought it necessary somehow to reconcile the absurd idea of negative-energy electrons with the fact that such monster particles were formally included in the theory's mathematical scheme. He was inclined to think that since the negative-energy solutions had a mathematical existence, they also had to be, not necessarily physically existing, but at least physically interpretable.[9] This reasoning, viz. what is in formal accordance with the laws of nature is also somehow to be endowed with physical reality, clearly is a version of the old principle of plenitude. Much later, Dirac recalled about this early phase of the positron theory: "I was reconciled to the fact, that the negative energy states could not be excluded from the mathematical theory, and so I thought, let us try to find a physical explanation for them" ([53],p.144). This attempt to endow the mathematical structure with a physical interpretation soon led to "perhaps the biggest jump of all the big jumps in physics of our century", to quote Heisenberg ([94],p.271).

The fact that an electron with negative energy moves in an external field as though it is positively charged, makes it tempting to suggest that the negative energy solutions are simply referring to ordinary protons. Dirac considered this possibility but only to reject it as unjustified: apart from the problem of difference in mass, a negative-energy electron cannot possibly have a physical existence, as this would lead to paradoxical consequences (cf. above). Dirac then flatly rejected the idea of a negative-energy electron (whether being identical to the proton or not) on empirical grounds: "No particles of this nature have ener been observed," is Dirac's plain reasoning ([26],p.362).

This is of course a perfectly sound argument, but it is not entirely consistent with Dirac's otherwise so rationalistic a philosophy. This, as well as other evidence to be mentioned in the present work, shows that Dirac's philosophy of mathematical reasoning was not strictly followed by himself. Although Dirac always stressed the primacy of mathematical reasoning and beauty at the expense of empirical physics, he was neither a full-blooded apriorist nor an anti-empiricist in any strict sense. When it came to the point, Dirac shared the basic conviction of all, or virtually all scientists, that ultimately one cannot make science disregarding experience.

So Dirac did never seriously regard electrons with negative energy as physically existing entities. Although positrons are sometimes described as particles with negative energy, this is mistaken and has nothing to do with Dirac's theory.[10] Dirac's rejection of negative-energy particles did not mean, however, that he also rejected the notion of negative-energy 'things', somehow predicted by the mathematical structure of his theory. Twelve years later, Dirac put the matter in this way:

> "Negative energies and probabilities should not be considered as nonsense. They are well-defined concepts mathematically, like a negative sum of money, since the equations which express the important properties of energies and probabilities can still be used when they are negative. Thus negative energies and probabilities should be considered simply as things which do not appear in experimental results" ([40],p.8).

In this formulation we may again recognize the principle of plenitude, although in a weak version: Whether the mathematics implies the material existence of ordinary entities or merely the abstract existence of 'things' is, after all, two very different situations.

As a way out of the dilemma, retaining the significance of the negative energy states without introducing negative-energy particles, Dirac suggested the strange picture of a world of negative energy states - the 'Dirac sea' - occupied by an infinite number of electrons governed by Pauli's exclusion principle. This sea, if completely filled, has no observable properties whatsoever, a feature emphasizing the abstract nature of Dirac's picture. The occasionally unoccupied states in this world, the 'holes', are endowed with positive energy and positive charge. Dirac suggested at first that the holes in the sea of negative energy were to be identified with protons.

Dirac's startling idea of a negative-energy world provided a new and strange picture of the vacuum, previously thought of as nothing but emptiness. Now the vacuum was conceived as an infinite, uniform distribution of ghost particles, a picture which brings to the mind the aether models in vogue in the latter part of the 19th century (cf.[12]). Many years later, Dirac proposed to revive the aether concept (see §5).

In the beginning of the thirties, Dirac's theory was not generally accepted and hardly regarded to be a respectable physical theory. In most quarters, Dirac's idea was rejected, partly because of mistakes about the anti-electrons, which were often taken to material particles with negative energy, and therefore unacceptable.[11] In his authoritative 1933 Handbuch survey of quantum mechanics, written shortly before Anderson announced his discovery of the positive electron, Pauli was very critical to Dirac's hole theory. The difficulty about the negative energy states cannot, Pauli wrote, "weder weggeleugnet noch in einfacher Weise behoben". Dirac's and Oppenheimer's ideas of anti-particles were rejected with the following argument:

> "Das tatsächliche Fehlen solcher Teilchen wird dann auf einen speziellen Anfangszustand zurückgeführt, bei dem eben nur die eine Teilchensorte vorhanden ist. Dies erscheint schon deshalb unbefriedigend, weil die Naturgesetze in dieser Theorie in bezug auf Elektronen und Antielektronen exakt symmetrisch sind. Sodann müssten jedoch (um die Erhaltungssätze von Energie und Impuls zu befriedigen mindestens zwei) γ-Strahl-Photonen sich von selbst in ein Elektron und ein Antielektron umsetzen können. Wir glauben also nicht, dass dieser Ausweg ernstlich in Betracht gezogen werden kann" ([138],p.246).[12]

Another objection raised against the hole theory was that Dirac's infinite sea of negative-energy electrons should produce an infinite electrical charge density and then an infinite field. This difficulty was pointed out publicly by Oppenheimer [136], and by Bohr in his correspondence with Dirac. Dirac tried to circumvent the difficulty by interpreting the charge density in Maxwell's equation as the departure from the 'normal state' where each state of positive energy is unoccupied and each state of negative energy is occupied. This normal state, or perfect vacuum, was assumed to produce no field. Later, Dirac considered the problem mathematically and set up a theory for the production of the electromagnetic field due to the negative-energy electron distribution [31']; this field he sought to make physically comprehensible, i.e. finite, by use of a complicated mathematical 'cut out' technique.

Dirac's optimism as regards the essential validity of general quantum mechanics, a theme often appearing in his publications, as well as his morale of piecemeal generalizations, was expressed in an exchange of letters with Bohr, about 1930. While Bohr was at that time deeply worried about the soundness of quantum mechanics, which he considered to be in a state of severe crisis because of its difficulties in the relativistic domain, Dirac was quite confident that quantum mechanics, if only appropriately generalized, would account for any problem that might arise. To Bohr he wrote: "I cannot see any reason for thinking that quantum mechanics has already reached the limits of its development. I think it will undergo a number of small changes ... and by these means most of the difficulties now confronting the theory will be removed. If any of the concepts now used .... are found to be incapable of having an exact meaning, one will have to replace them by something a little more general, rather than make some drastic alteration in the whole theory."[13] As we soon shall see, Dirac was not able to maintain his optimism during the thirties.

Dirac's relativistic theory of electrons, positive and negative, has not aroused much interest from the side of philosophers of science. Apart from Bachelard, already mentioned, one of the few exceptions was due to Henry Margenau who, already in 1935, subjected Dirac's theory to a methodological analysis [174]. The constructs of explanation, appearing in Dirac's theory, are purely abstract, that is, without any visualizable counterpart in nature. This, in itself, is not disqualifying to the value of the theory, Margenau points out. But the theory introduces constructs (such as the 'sea' of negative-energy electrons) to which, by definition, no counterpart even in the form of data correspond. And hence the theory sins against the requirement of simplicity, in the same way as did the mechanical aether models or other hypotheses involving hidden mechanisms. Margenau applied his methodological criticism also to the then new neutrino hypothesis, considered to sin against the principle of simplicity even more violently. For the neutrino hypothesis involves not only the construct of neutrinos, supposed to be practically beond empirical tests; it also admits "the existence of higly paradoxical entities which could be called "anti-neutrinos""(ibid.,p.88), and is therefore double methodologically unsound.

With the advantage of hindsight it is instructive to see how Margenau's early methodologicla analysis, explicitly designed to be an "attempt at crystallizing what the physicist is doing," and "not to be regarded as a program to which physical activities must always conform" (ibid.,p.81), nevertheless leads to the rejection of physical notions alone from methodological considerations. What Margenau in 1935 called "monstrosities" and "highly paradoxical entities", anti-neutrinos, are in fact familiar objects for todays physicists. Also it is characteristic that Margenau, writing after the discovery of the positron, judged Dirac's theory as methodologically better than Fermi's. The first one was regarded to be "incomplete in its interpretation" but not "altogether misleading", such as was Fermi's neutrino theory. One might wonder what Margenau's methodological judgment would have been if made in 1931, before the positron, or in 1957, after the neutrino was experimentally confirmed. If there is a morale to draw from this story, it would be one rather agreeing with Feyerabend's against-method philosophy, I think.

In Dirac's original identification of holes with protons, various motives played a role. At first Dirac was inclined to consider the holes as being positive electrons. According to the mathematical structure of the theory, one should expect a perfect symmetry so that the holes would have the same mass as the electron. If Dirac had kept faith in his 'principle of mathematical reasoning', he would at once have suggested the idea of a positively charged electron. However, other factors conflicted with such a proposal. Firstly, Dirac was enough of a physicist to prefer his theory to be based, if possible, on known entities. In 1930, the proton was the only known positively charged elementary particle. Secondly, the idea of identifying protons with vacant negative-energy states was highly attractive by standards of principle, as in this way all the then known elementary particles could be reduced to one fundamental particle.[14] To Dirac this 'principle of unity' no doubt played an effective role. In September 1930 he emphasized the point:

> "It has always been the dream of philosophers to have all matter build up from one fundamental kind of particle, so that it is not altogether satisfactory to have two in our theory, the electron and the proton. There are, however, reasons for believing that the electron end proton are really not independent, but are just two manifestations of one elementary kind of particle" ([28],p.605).

If the proton was the hole, one would expect that a positive-energy electron might occasionally make a quantum transition to fill the hole, under which circumstances the two particles would annihilate under emission of electromagnetic radiation ($p^+ + e^- \rightarrow 2\gamma$). This process had earlier been considered, first by Eddington in 1917 who suggested it as a source of stellar energy.[15] Dirac commented on the hypothetical process: "There appears to be no reason why such processes should not actually occur somewhere in the world. They would be consistent with all the general laws of Nature ...."([28].p.606). This statement is very typical for Dirac's philosophy of physics. In it, we meet again the principle of plenitude: what is consistent with "the general laws of Nature", probably is also realized in nature.

It should be remarked that Dirac, in 1931, also ventured a prediction of anti-protons, being negatively charged protons. These particles were only discovered, or rather manufactured, in 1955, 24 years after their prediction. In his Nobel Lecture [31], Dirac went one step further; from the view of complete symmetry between positive and negative electrical charge he speculated about anti-matter, entire stars being build up from positrons and anti-protons. Such anti-matter stars have later played a role in certain cosmological models (the Alfvén-Klein model).

Shortly after Dirac's hole theory was proposed, efforts were taken to examine it theoretically. Oppenheimer was able to calculate the mean lifetime for ordinary matter on Dirac's proton-electron theory to be of the order of 10^{-10} seconds, which is, as he said, "absurdly short" ([135],p.943). Similar calculations were also made by Dirac, who considered the transition probability for the process $p^+ + e^- \rightarrow 2\gamma$ [27]. His result was inconclusive but rather indicated a much too large value to agree with the known stability of the matter. Despite these evidences, Dirac tried to rescue his hypothesis, and then the "dream of philosophers", by arguing that the unknown interaction between the electron and the proton would possibly account for the gross discrepancy. It was only after Oppenheimer had brought more conclusive arguments against Dirac's assumption, and after Weyl had shown that the holes must necessarily have the same mass as ordinary electrons [167], that Dirac was forced to leave his unitary view. The proton was now exchanged with the positive electron, experimentally being an unknown particle.

In his various recollections, Dirac has repeatedly pointed out that his disinclination to postulate positively charged electrons was rooted in lack of boldness to rely on the mathematical results of his wave equation and disregard the restrictions put by current empirical knowledge. Had he only been faithful to his mathematical reasoning, and not been led astray by what was known experimentally, then he should immediately have postulated the existence of positrons. Indeed, this was what Weyl did, and Dirac attributes this success to Weyl exactly because he was not a physically minded physicist:

> "... Weyl was a mathematician. He was not a physicist at all. He was just concerned with the mathematical consequences of an idea, working out what can be deduced from the various symmetries. And this mathematical approach led directly to the conclusion that the holes would have to have the same mass as the electrons. Weyl .... did not make any comments on the physical implications of his assertion. Perhaps he did not really care what the physical implications were. He was just concerned with achieving consistent mathematics." ([51],p.55)

Curiously, it was in following another suggestion of Weyl, from 1929,[16] that Dirac first tried to construe the negative energy solutions as protons.

During the thirties, Dirac became increasingly worried about the state of affairs in quantum theory, especially in relation to the intriguing problems arising from the infinite number of negative-energy electrons and from the appearance of divergent integrals in quantum electrodynamics. These problems made the theory aesthetically unattractive to Dirac. They led, Dirac stated in the 1941 Bakerian Lecture, "to such complicated mathematics that one cannot solve even the simplest problems accurately, but must resort to crude and unreliable appproximations. Such a theory is a most inconvenient one to have to work with, and on general philosophical grounds one feels that it must be wrong" ([40],p.9). The growing dissatisfaction made Dirac give up his usual optimistic attitude and conclude that a drastic change in the fundamentals of quantum mechanics was necessary. When Shankland in 1936 announced results of photon scattering experiments which seemed to imply non-conservation of energy for individual atomic processes, Dirac used this opportunity to a showdown with the "so-called quantum electrodynamics" [33]. He argued that Shankland's measurements necessitated that the current relativistic quantum field theory, which involved conservation of energy and momentum, had to be replaced by a new theory of the Bohr-Kramers-Slater type. This short-lived but influential theory, originally proposed in 1924 as a rather desperate answer to the growing crisis in quantum theory, gave only a statistical conservation of energy, not conservation for individual processes (cf. [101],pp.181-188). Willing to sacrifice the troublesome relativistic theory, Dirac wanted to retain the general, non-relativistic theory and to take it as the starting point for a new and better generalization, based on non-conservation in relativistic processes: "The present quantum mechanics, with its conservation of energy and momentum, forms a satisfactory theory only when applied non-relativistically, to problems involving small velocities; and loses most of its generality and beauty when one attempts to make it relativistic" ([33],p.298). Dirac ended his short paper, "... we may give it [quantum electrodynamics] up without regrets - in fact, on account of its extreme complexity, most physicists will be very glad to see the end of it" (ibid.,p.299). In this judgment, Dirac probably gave a correct picture of the situation. Throughout the thirties, most physicists thought that quantum field theory, though applicable to a number of problems, did not really make sense and that some radical new ideas had to be added (cf. [162]).

It is hard to believe that Dirac accepted Shankland's results simply as an experimental evidence with which the theory had to comply. Such as empiricist attitude would be in gross contrast to his general ideas about the relationship between theory and experiment in physics (see §9). Rather, he used Shankland's experiment as an additional argument in support of his theoretically based disbelief in quantum electrodynamics. For Dirac, the really important thing about Shankland's experiment was that it called attention to the existing inadequacies in relativistic quantum theory.

Dirac's doubt as to the "so-called quantum electrodynamics" naturally also covered the theory of β-decay, based upon energy-conserving relativistic quantum mechanics. As well known, Pauli had in 1930 tentatively suggested the existence of mass-less particles,[17] by Fermi later dubbed 'neutrinos', in β-processes in order to reconcile the experimental data with the principle of energy conservation as well with conservation of spin and statistics. Pauli's outlandish neutrino hypothesis was developed quantitatively by Fermi in his 1934 theory of β-decay, which soon proved highly successful. In theories of the Bohr-Kramers-Slater type, however, there is no need to postulate neutrinos, in which Dirac did not believe and considered as introduced *ad hoc*: "... a new unobservable particle, the neutrino, has been specially postulated by some investigators, in an attempt to preserve conservation of energy by assuming the unobservable particle to carry off the balance"(ibid.).

Dirac's unsympathetic attitude to the neutrino may seem surprising in regard to his general philosophy of physics. Indeed, Pauli's neutrino postulate may seem to be based on a similar reasoning which made Dirac postulate the anti-electron and the magnetic monopole. Clearly, Dirac would not accept the neutrino simply on the ground that it was a conceivable entity which rescued some generally believed conservation principles. In Dirac's version of the principle of plenitude, he wanted a mathematically consistent reason in order to postulate physical entities. They should, so to speak, grow out of the equations' mathematical structure. It can hardly, however, have escaped Dirac's attention that in fact there were also purely theoretical grounds for considering mass-less neutral particles of spin ½ in β-processes. Pauli had constructed a variant of Dirac's equation which described his hypothetical particle and shown that the neutrino was, in this sense, allowed by relativity and quantum theory. In 1932, Oppenheimer and Carlson, examining Pauli's neutrino hypothesis, remarked that it was put forward "on the further ground that such a particle could be described by a wave function which satisfies all the requirements of quantum mechanics and relativity".[18] From these theoretical evidences to infer the existence of a neutrino, would certainly be an argument in complete agreement with Dirac's way of thinking. So, for Dirac in 1936, two aesthetic principles, both of significance in his general reasoning, gave opposite answers to a specific problem: From the principle of plenitude, in Dirac's key, he should have welcomed the neutrino; from the principle of mathematical beauty, applied to the unappealing mathematical structure of the quantum field theory, he was however led to discard it.

In 1936 Dirac developed his 1928 relativistic wave equation a step further so as to include also particles of spin larger than a half [32]. Typically, Dirac was not pressed to do so by new experimental evidences. In the beginning of 1936, only the electron (positive and negative), the proton and the neutron were recognized as elementary particles, each of them carrying a spin of one half. To be true, Yukawa had in 1934 suggested nuclear forces to be due to a new field whose quantum was estimated to be about 200 electron masses. In 1935-36, Yukawa's prediction of the meson was examined by Japanese physicists (Mituo and others) who argued that the spin of the meson was one unit (the true value is zero). In 1937, Yukawa's particle was erroneously identified with the 'mesotron', detected in the cosmic radiation. It is unlikely, however, that Dirac's extension of his theory was influenced by the prediction of the meson. In 1936, very few people accepted Yukawa's new particle, himself admitting it to be based on "arguments .... of merely speculative character" ([173],p.57). Mituo's calculation of the meson's spin was, moreover, not known in Europe [131].

So it is, in effect, almost certain that Dirac's 1936 theory was not motivated by this early phase of meson theory. As far as Dirac was concerned, no evidence of particles with spin larger than a half asked for theoretical treatment. All the same, Dirac reformulated his original linear wave equation in a more general way, and showed that it was then able to describe also particles with spin larger than a half, whether with zero mass or not. Dirac justified what might appear to be merely a mathematical exercise, as follows: "It is desirable to have the equations ready for a possible future discovery af an elementary particle with a spin greater than a half, or for approximate application to composite particles" ([32],p.448). And in a characteristic spirit it was added: "Further, the underlying theory is of considerable mathematical interest."

Let it finally be mentioned that in his later days, Dirac has resumed his early interest in relativistic wave equations and has proposed a theory which allow only for positive energy solutions [50]. The new wave equation is, of course, relativistic invariant, and it bears a formal similarity to the celebrated 1928 equation. The particles it describes are not, however, electrons but some new entities, not known to exist in nature. Although modern elementary particle physics may be said to owe its origin to the works of Dirac, he has always been remarkably silent about problems in elementary particle and nuclear physics, the two most fashionable and fast-growing branches of post-war physics. Apparently this kind of physics had no appeal to

Dirac, who preferred a more abstract and general line of thinking. I know of only one occasion where Dirac has indicated his view about elementary particles. This was following a talk by Heisenberg in which he expressed his deep dislike of the Democritean philosophy underlying the physicists' search for the ultimate constituents of matter, including the 'prejudice' in favour of quarks. Dirac agreed with Heisenberg's point of view about elementary particles that a concept really doesn't exist, but added that he was inclined to consider the electron as elementary. "It may be that I am prejudiced because that I have had success with the electron and no success with other particles," he said ([94],p.274).

At an early state, Dirac considered the success of his relativistic electron theory as rooted in the more or less explicit philosophy behind the taken approach. This view was expressed in his preface to <u>Principles of Quantum Mechanics</u> and, in greater details, in May 1931, introducing one of Dirac's most important papers. This introduction, a most interesting document of a young working physicist's philosophy of science, deserves to be quoted at length:

> "The steady progress of physics requires for its theoretical formulation a mathematics that gets continually more advanced. This is only natural and to be expected. What, however, was not expected by the scientific workers of the last century was the particular form that the line of advancement of the mathematics would take, namely, it was expected that the mathematics would get more and more complicated, but would rest on a permanent basis of axioms and definitions, while actually the modern physical developments have required a mathematics that continually shifts its foundations and gets more abstract. Non-euclidean geometry and non-commutative algebra, which were at one time considered to be purely fictions of the mind and pastimes for logical thinkers, have now been found to be very necessary for the description of general facts of the physical world. It seems likely that this process of increasing abstraction will continue in the future and that advance in physics is to be associated with a continual modification and generalisation of the axioms at the base of the mathematics rather than with a logical development of any one mathematical scheme on a fixed

foundation.

There are at present fundamental problems in theoretical physics awaiting solution, e.g., the relativistic formulation of quantum mechanics and the nature of atomic nuclei (to be followed by more difficult ones such as the problem of life [sic!]), the solution of which problems will presumably require a more drastic revision of our fundamental concepts than any that have gone before. Quite likely these changes will be so great that it will be beyond the power of human intelligence to get the necessary new ideas by direct attempts to formulate the experimental data in mathematical terms. The theoretical worker in the future will therefore have to proceed in a more indirect way. The most powerful method of advance that can be suggested at present is to employ all the ressources of pure mathematics in attempts to perfect and generalise the mathematical formalism that forms the existing basis of theoretical physics, and after each success in this direction, to try to interpret the new mathematical features in terms of physical entities (by a process like Eddington's Principle of Identification)." ([30],pp.60-61).

§ 5. The Theory of Magnetic Monopoles.

In 1931 Dirac used his 'general scheme of advance' to put forward a quantum theory of magnetic charges [30]. This theory did not share the success of his electron theory, in the sense that it was not experimentally verified despite numerous attempts and even one claim. On the whole, Dirac's theory of magnetic poles has, though widely debated, always been considered to be a little speculative and not entirely convincing. This incidentally, was also the general reaction to the hole theory before the discovery of the positron. Dirac, however, has remained faithful to the ideas of his original theory. In 1948 he developed the theory further [41], and also at later occasions Dirac has taken it up. But even if the fate of Dirac's magnetic theory has been very different from the celebrated Nobel Prize winning electron theory, otherwise the two theories show a nice intrinsic harmony. In its structure and philosophy, Dirac's theory of magnetic poles accords well with his slightly earlier theory of anti-electrons.

In classical electromagnetic theory, such as formulated by Maxwell and Lorentz, electricity and magnetism appear in an integrated but not entirely symmetric way. In particular the units of electrical charge have no counterpart in magnetic charges, for which there is no experimental evidence. Maxwell originally discussed 'magnetic matter' but only, he stressed, in "a purely mathematical sense" ([120],art.380). If magnetic matter or 'fluids' should account for the observed phenomena, one had to introduce as an extra axiom that such matter is confined within the molecules so that macroscopic bodies containing an excess of one magnetic fluid are avoided. Maxwell ended up with the standard view of magnetic poles as being the ends of long thin magnets, and dismissed proper magnetic monopoles simply because there was no experimental evidence for them. Still, there is nothing in the Maxwell-Lorentz theory that precludes that such entities might possibly exist. In fact, the motion of electrical particles in fields generated by magnetic monopoles was discussed before Dirac was even thought of. Poincaré [140] and J.J.Thomson [159] were among the physicists who examined the motion of magnetic monopoles in a classical context. Another author who dealt with the question of isolated magnetic poles was, incidentally, Friedrich Engels.[19] It is only in the more modern formulations of electrodynamics, where the electromagnetic potentials are crucial and where the equations of motion are expressed in Hamiltonian form, that there is no room for magnetic monopoles. Since a canonical formulation of the electromagnetic field turned out to be necessary for a quantum description of the field and its interaction with particles, it was thought that the existence of magnetic charges was incompatible with quantum mechanics (see also appendix II).

Dirac objected to this conventional thinking and shoved that magnetic monopoles can perfectly well exist consistently with quantum mechanics if it is only slightly generalized. Just as the hole theory aimed to introduce a symmetry between electrons and protons, the 1931 theory aimed to introduce a symmetry between electricity and magnetism. But just as the electron-proton symmetry is mot conplete, neither is the symmetry between electricity and magnetism, Dirac explained.

From 1926 onwards, Dirac has taken the validity of general quantum mechanics as a basis for most of his research in the qantum domain. The confidence in the basic quantum mechanical formalism was decisive for Dirac's theories of 1928 and 1930, and also in 1931 he emphasized that his new theory was firmly based in the fundamental theorems of quantum mechanics.

> "The present formalism of quantum mechanics, when developed naturally without the imposition of arbitrary restrictions, leads inevitably to wave equations whose only physical interpretation is the motion of an electron in the field of a single pole. This new development requires <u>no change whatever</u> in the formalism when expressed in terms of abstract symbols denoting states and observables, but is merely a generalization

of the possibilities of these abstract symbols by wave functions and matrices" ([30],p.71).

In accordance with his mathematical programme, his 'general scheme of advance', Dirac did not start out with experimental puzzles but with mathematical deductions to which a physical meaning was afterwards assigned. Dirac showed that the unphysical phase factor in the wave function $\Psi=\psi e^{i\gamma}$ may be related to singularities in the electromagnetic field, pictured as discrete Faraday lines of force (see appendix III). These singularities were interpreted as electrical and magnetical point charges. Elaborating on these ideas, Dirac argued that quantum mechanics allows formally for magnetic monopoles. And further, <u>if</u> such entities are postulated, then it follows necessarily from the logic of quantum mechanics that their magnetic strenght is quantized according to

$$\mu = n \cdot \mu_0 \qquad \text{with} \qquad \mu_0 = \frac{\hbar c}{2e} \qquad \text{(a)}$$

where n is an integral number.

Apart from the desire to bring more symmetry into physics, Dirac was no doubt motivated also by an ambition to explain the atomicity of the electrical charge. Not merely accepting the existence of a smallest electrical charge as a matter of experience, Dirac asked about its deeper reason. He seems to have hoped for a purely quantum condition, expressing the magnetic pole in analogy with the electric one achieved from the fine-structure constant

$$\alpha = \frac{e^2}{\hbar c} = \frac{1}{137} \qquad \text{or} \qquad e = \sqrt{\alpha \hbar c} \qquad \text{(b)}$$

What Dirac deduced was, instead of, a reciprocal relation between magnetic and electric charges, a result he found "rather disappointing". But he also considered it to be a strong argument in favour of the existence of magnetic monopoles. Because such particles will, by virtue of (a), require all electrical charges to be quantized in units of $\hbar c/2\mu_0$.

As Dirac stated:

> " The quantization of electricity is one of the most fundamental and striking features of atomic physics, and there seems to be no explanation for it apart from the theory of poles. This provides some grounds for believing in the existence of these poles" ([41], p. 817).

Dirac thought erroneously that his theory was the first attempt to explain the atomicity of electrical charge from a fundamental theory.[2o]

What Dirac showed was, that magnetic monopoles can exist, according to quantum mechanics. But do they really exist?

In accordance with his stated methodology of mathematical reasoning, Dirac was inclined to think that since there is no theoretical reason against the existence of magnetic monopoles, then they would probably also exist in nature. "Under these circumstances one would be surprised if Nature had made no use of it," Dirac said ([3o], p. 71), virtually echoing earlier statements from his theory of electrons.[21] This is a rather precise statement of the principle of plenitude, also operative in Dirac's positron theory. The principle of plenitude works, however, somewhat differently in the two cases. While in the latter, anti-electrons are not only formally allowed, but are also demanded - as the theory would be incomplete without them - in the first theory monopoles are predicted solely by a negative argument: they are not precluded by the theory. Quantum electrodynamics is neither more, nor less, complete or consistent whether the monopoles are postulated or not; but of course the theory attains a more symmetrical form. In general one should take care to distinguish between phenomena explicitly predicted ('demanded') by theory and phenomena which are just not precluded by theory.

Due to its inexactitude, the principle of plenitude has, in the course of history, been used as an argument for the physical existence of almost everything, from zoophytes to monopoles. At times, indeed, reasoning based on the principle of plenitude seems to amount to little more than asking "why not?" In the eighteenth century, for example, the principle was used to support the claimed observations of mermaids and sea-men by a priori arguments. As the notion of mermaids was neither intrinsically contradictory nor colliding with current biological laws, such creatures were, in the light of the principle of plenitude, assumed to exist in nature ([117], p. 271). In its essence, Dirac's line of reasoning when conjecturing the existence of magnetic monopoles did not differ from the eighteenth century arguments in favour of the existence of mermaids. Although both mermaids and monopoles have been claimed observed from time to another, neither of them have found widespread acceptance among scientists.

The reception of Dirac's audacious theory was remarkable because of its silence and lack of interest. In the period 1931-39, only five scientific papers dealt with magnetic monopoles,[22] excluding Dirac's own contribution. Apparently Jordan was the only one who took an active interest in Dirac's theory; of the five papers on monopoles, two were written by Jordan and one by his assistant B.O. Grönblom. The scant interest devoted to Dirac's theory was, moreover, restricted to a purely technical level. Igor Tamm solved the differential equation governing the motion of an electron around a monopole [158], and Jordan proved that the Dirac condition for monopole strength may also be derived from quantum electrodynamics [1o5]. In the thirties, nobody commented on Dirac's suggestion about the physical existence of monopoles, and neither was there any experimental search for them.

On remarkable exception was Felix Ehrenhaft, the Austrian physicist who unhappily devoted his scientific life to the discovery of subelectronic charges [96]. In a number of experiments, Ehrenhaft found charges which were fractions of the elementary charge and insisted on the significance of his results. Other physicists, however, flatly refused to accept his claim of subelectrons. Despite his important contributions to physics (photophoresis and Brownian motion), he was soon regarded as nothing but a crank. Around the mid-thirties, after Dirac had put forward his theory of magnetic monopoles, Ehrenhaft claimed to have discovered monopoles and continually wrote to Dirac about it.[23] He tried to obtain support from Dirac who, however, concluded that Ehrenhaft's experiments could not be interpreted as evidence for his theory and he was not able to support the outcasted Ehrenhaft. In the thirties and forties, Ehrenhaft was virtually banned from publication in the recognized physics journals. In 1945, he was allowed to report briefly on his experiments which he interpreted in terms of particles with a single magnetic pole [68]. In 1948 Dirac referred to this report, but did not accept it as evidence for the magnetic monopoles belonging to his theory ([41], p. 817).

The fact that magnetic monopoles have never been detected seems not to bother Dirac overly. To Dirac the really important thing is not whether such particles are experimentally confirmed or not but that they have, at least, a theoretical existence (or, what is practically the same, that they exist "somewhere in the universe"). Contingency, rather than empirical existence, was Dirac's prime interest. Still, Dirac was concerned with physics, not metaphysics, and in 1931 he ended his article with a few remarks about the possible detection of the magnetic poles, much like the way that he had earlier commented on the possible experimental verification of his "hole" theory. Dirac suggested that the lack of observation of magnetic monopoles was rooted in the large numerical value of μ_0, about 7o times the electrical unit charge (see (a) and (b)). This requires enormous energies to produce and separate pairs of monopoles.

Reading Dirac's papers, one gets the impression that he never cared too much about experimental confirmation. In 1975 he payed, indeed, a tribute to empirical philosophy: "Whether they [magnetic monopoles] exist or not can only be decided by experiment" ([56], p. 46). It was, in 1975 as in 1931, obvious to Dirac that physical existence can only be truly decided by empirical means. Dirac shared, after all, the unconscious empiricism of most physicists. Yet, he seems to have been quite satisfied with the internal theoretical power of his magnetic theory, whether confirmed or not. When P.B. Price and co-workers in 1975 reported that they had detected a single event of a magnetic monopole in the atmosphere,[24] one would expect that Dirac would enthousiastically hail this discovery, long waited for. But he welcomed it only rather indifferently. While Price's discovery was considered much sensational in the press and by most physicists, and gave rise to heated discussions, Dirac sufficed to state that there was probably a fifty-fifty chance that Price was right; and soon after he told Price that he did not believe in his discovery.[25]

This rather unusual reaction indicates a relative lack of interest in empirical verification and a preoccupation with the theory's intrinsic qualities. These characteristic features in Dirac's psychology were also manifest in 1932 when Anderson first detected the positive electron and thereby confirmed Dirac's controversial theory of "holes". In an interview from 1963, T.S. Kuhn asked whether this discovery generated great immediate excitement and satisfaction. Dirac answered: "I don't think it generated so much satisfaction as getting the equations to fit."[26]

Some years after Dirac's second contribution to monopole theory, he proposed to revive the aether, that old favourite medium of British physicists. Though without technical connection to monopoles, Dirac's idea of an aether shows a certain methodological similarity to his works on monopoles which is the reason that I mention it here.

Dirac's unorthodox attempt to reintroduce the aether in physics was based in his continuous dislike of the state of art in quantum field theory. As mentioned in the previous section, this dislike was predominant also in the thirties, then shared by a majority of physicists. After the war, however, quantum field theory gained a momentous break-through as Schwinger, Dyson, Feynman, Tomonaga and others managed to develop techniques for removing the infinities which had haunted the earlier phases of the theory (see [162]). In the late forties, quantum field theory was developed into a highly successful theory, generally recognized to be the relativistic quantum theory. Dirac, however, remained sceptical and was not convinced by the new theory. It lacked, he felt, the two qualities which characterize a really satisfactory theory, generality and mathematical beauty.

Dirac's judgement of quantum field theory, anno 1954, was essentially the same as 15 years earlier. His main objection was that, "the present quantum field theory is complicated and ugly. It has none of the simplicity and beauty which are characteristic of a good physical theory. These qualities occur to a marked extent in relativistic mechanics alone, or in quantum mechanics alone, but disappear with our present methods of combining the two" ([44], p. 145). In 1936, Dirac proposed to retain beauty by giving up energy conservation; now, 15 years later, he argued for a no less controversial idea, the aether, in order to reach the same goal.

Since the late thirties, Dirac had attempted to formulate a new classical theory of electrodynamics, amenable to a more satisfactory quantization procedure. This programme started in 1938 [37] and was further developed after the war [43]. It was well known that the problem of infinities in quantum electrodynamics was essentially carried over from the classical Lorentz theory, in which the self-energy of the point electron is infinite. Therefore Dirac applied the strategy that the problem of infinities should first be solved in classical electrodynamics and quantization only come afterwards, impressed on the improved classical theory. Possibly, Dirac stated his idea, "the troubles of the present quantum electrodynamics should be ascribed primarily .... not to a fault in the general principles of quantization, but to our working from a wrong classical theory" ([43], p. 291). "The theory .... is put forward as a basis for a passage to a quantum theory of electrons" (ibid., p. 296), the aim was declared. In Dirac's reinterpretation of the Maxwell-Lorentz equations, the electric charges were not introduced as new variables, but were described by "superfluous" variables appearing in the theory without charges. These variables, corresponding to the choice of a gauge in the Maxwell-Lorentz theory, were, Dirac stated, "to a certain extent, <u>at our disposal</u>, and .... they can be made to serve in the description of electrons, instead of remaining physically meaningless" (ibid., p. 293). Notice the close resemblance of this argument, again a kind of plenitude reasoning, with the earlier argument for the negative energies, as expressed in e.g. 1942 (se §3). Dirac thus tried to construct a scheme of electrodynamics in which electric charges appeared only as secondary entities. Or, to be more correct, his theory involved only the ratio e/m, not e and m separately. The existence of an electronic charge was thought not to be a property of classical electrons, but to be a quantum effect. Despite much work and ingenuity directed to this programme, it did not succeed. That is, it never reached quantization and most physicists did not accept it.

A peculiar feature in Dirac's new electrodynamics, such as it was developed in 1951, was that the electromagnetic potentials were interpreted as velocities.[27] In vacuum, the particular velocity was taken to be the velocity with which a small charge would move, if introduced. On this new conception, Dirac proposed to revive the aether, thought to be ruled out of physics by the theory of relativity in the beginning of the century [42]. Dirac argued that on his new electrodynamics an aether could well be conceived, and that without violating the basic laws of relativity and quantum mechanics. The velocity of Dirac's aether was taken to be the afore-mentioned vacuum velocity associated with the electromagnetic potentials. As in the 193o hole theory, Dirac viewed the vacuum as being different from mere emptiness. The vacuum, he noticed with satisfaction, "is no longer a trivial state, but needs elaborate mathematics for its description" ([42], p. 9o6). Also the demonstration that the aether is reconcilable with basic physical laws and by that reason is likely to exist, runs parallel to the 1931 argument for the monopole. Dirac ended his short note in 1951 with the words: "It is natural to regard it [the Dirac vacuum velocity] as the velocity of some real physical thing. Thus with the new theory of electrodynamics we are rather forced to have an aether" (ibid., p. 9o7).

Dirac realized, of course, that the classical aether concept, involving a definite aether velocity and hence giving preference to one direction in space-time, was incompatible with the principle of relativity. But what does the matter look like, Dirac asked in 1954, if viewed in the light of quantum mechanics? In this case one would think that the aether, imagined to be a very light form of matter, would be strongly affected by the principle of indeterminacy. If so, the velocity of the aether would not be definite but subject to a quantum mechanical probability law, and therefore conceivable within the principle of relativity. Together with the aether, Dirac now suggested to bring in absolute time, similarly adjusted to the requirements of quantum mechanics. In this way, it was conjectured, the unattractive complexities of quantum field theory would be avoided and the inherent simplicity of non-relativistic quantum mechanics restored into the relativistic domain. How this should be carried through was not explained as Dirac was not able to develop his ideas about a new aether and a new absolute time into a mathematical theory. As to whether the aether really exists, Dirac partly left the matter to the experimentalists: If aetherbased theories would turn out to be empirically better than aetherless theories, then the aether would "exist". Dirac concluded:

> "I would like to emphasize that the foregoing discussion does not prove the existence of an aether or of absolute time. It merely shows that these concepts are not inconsistent with relativity, when one applies them in a setting of quantum mechanics, and so there is no immediate reason for rejecting them. Whether nature has actually made use of them or not is another question" ([44], p. 146).

Despite of his cautious choice of words, it seems that Dirac considered the aether as more than just a working hypothesis of possible heuristic value. Being consistent with quantum mechanics and relativity, the aether was conceived as "really existing" in the same vague sense as was the magnetic monopole. In the last analysis, however, this existence would be revealed only through its contribution to a mathematically simpler and more beautiful physical theory.

It may be remarked that Dirac's dissatisfaction with the situation in quantum field theory, a branch which was pioneered by himself in 1927 and to which he has contributed as much as any other physicist, has never ceased. In 1964, in lectures given at Yeshiva University, Dirac described the usual form of quantum field theory as being "in rather a mess", as "a stopgab, without any lasting future" ([47], pp. V and 2).

§ 6. MONOPOLES AND TACHYONS IN THE THEORY OF SCIENCE.

Let us now, for a moment, digress from the history of Dirac's physics and inquire into some general aspects of the meta-particles occuring in modern physics.

The arguments for the existence of the magnetic monopole are clearly rooted in two aesthetic principles: A principle of plenitude and a symmetry principle, the latter stating that the electromagnetic equations ought to be symmetric in electricity and magnetism. From these principles it follows that a symmetric theory, including monopoles, is to be preferred to an asymmetric theory without monopoles.

As earlier mentioned, classical electrodynamic theory is formally consistent with magnetic monopoles. It can be shown, however, that this peaceful coexistence breaks down for certain formulations of electromagnetic theory such as electromagnetism framed canonically, i.e. by means of Hamilton-Jacobi formulation. Also it turns out that it is not possible to give an action principle for the classical electron-monopole electromagnetic field unless an extra condition, not derivable from the action principle, is assumed.[28] As was pointed out in §3, fundamental laws in physics are thought always to be formulable by action principles. At least one physicist objected to the 1975 claim of discovery because he thought that particles, not obeying pure action principles, would be intolerable. His argument is interesting because it shows (1) that the aesthetic factor is clearly recognized, and (2) how two aesthetic criteria, both derived from the quest of beauty, may contradict each other in a specific case. Symmetry and plenitude versus action principle, which should be ranked highest?

The mentioned physicist, David Rosenbaum, objected to Price's claim of having discovered the Dirac monopole in the following way:

> "The arguments for magnetic monopoles are essentially aesthetic. The fact that the electric-charge-magnetic-monopole system, if it existed, would be the only classical system whose dynamical equations could not be derived from an action principle destroys

any aesthetic advantage for me, and thus any attractiveness to the concepts." [15o]

The fact that electrodynamics with magnetic monopoles is difficult to express satisfactorily in terms of an action principle implies, furthermore, that monopoles are impossible within the framework of direct-action electrodynamics [16o]. Only in field theories of the Maxwell type are monopoles permitted. This has led some physicists to suggest that empirical disproof of magnetic monopoles would imply an abandonment of electromagnetic fields and instead of support ideas of action-at-a-distance. It is only if fields are given the same ontological status as particles, that the apparent absence of monopoles in nature becomes a mystery.

Among the multitude of particles, or meta-particles, occuring in modern physics, some show a striking analogy to Dirac's monopoles. This is the case with the so-called tachyons, that is hypothetical particles which move with a speed faster than light.

Tachyonic particles have always been considered as conflicting with the principle of relativity, based upon the postulate of the velocity of light as an upper limit. In pre-relativity physics faster-than-light particles were discussed by Arnold Sommerfeld in 19o4 [156], and in greater details in 19o5 [157], and also by some other authors. Sommerfeld's investigations, which built on the classical electron theory, showed that particles accelerated beyond the light barrier would behave in a manifestly absurd way. Not only would they accelerate upon loss of energy, but if ascribed a mass as force divided with acceleration, it would be negative. Because of these absurdities, it was hard to believe that superluminal electrons existed in nature. All the same, tachyons were not really precluded by classical theory. Sommerfeld, for instance, did not consider the idea of material bodies moving with a velocity faster than light as an unphysical concept in itself. Cf. [147].

While faster-than-light particles could at least be sensibly discussed in pre-relativity physics, the theory of relativity exiled such particles from physics to the airy realms of metaphysics. In the sixties, however, physicists reexamined the matter from the viewpoint of special relativity and demonstrated that this theory does not, after all, preclude superluminal particles, tachyons.[29] Only they are bound to remain superluminal. Crossing the light barrier definitely violates the principle of relativity. Also it is possible to describe tachyons, if assumed to be spinless, by the formalism of relativistic quantum field theory. This was shown by Feinberg [73] who also coined the term tachyon, referring to the Greek word "tachys", meaning swift. So tachyons do not involve logical inconsistencies, and may be reconciled with the fundamentals of physical theory, relativity and quantum mechanics.

From the fact that tachyons are not precluded by theory, they were then assumed to exist in nature. This argument, derived by the principle of plenitude, was fully recognized by the involved physicists. Bilaniuk and Sudarshan, who first predicted/invented the tachyons, wrote:

> "There is an unwritten precept in modern physics, often facetiously referred to as Gell-Mann's totalitarian principle, which states that in physics "anything which is not prohibited is compulsory." Guided by this sort of argument we have made a number of remarkable discoveries, from neutrinos to radio galaxies." ([5], p. 44).

The close agreement between monopoles and tachyons, as far as their meta-scientific status is concerned, is further shown by the following statement, excerpted from a 1963 article on magnetic monopoles:

> "One of the elementary rules of nature is that, in the absence of law prohibiting an event or phenomenon it is bound to occur with some degree of probability. To put it simply and crudely: Anything that can happen does happen. Hence physicists must assume that the magnetic monopole exists unless they can find a law barring its existence." ([78], p. 122)

Here we have, in clear language, the modern principle of plenitude, much in the sense that it was applied by Dirac. The agreement with Robinet's 18th century formulation (see §3) is striking.

The similarity with Dirac's 1931 work is distinct also in Bilaniuk and Sudarshan's concluding obligatory tribute to empiricism: "Although we think that superluminal particles do exist, the only unequivocal way to ascertain that they do is by actually detecting them." ([5], p. 51). Ever since their prediction in the late sixties, tachyons have been sought for experimentally, but hitherto in vain. Curiously, tachyons were reported detected in 1974 [18], one year before Price's monopole claim. However, the claim turned out to be too hasty and was soon withdrawn.

The latest development in the hunt for new particles, assumed to exist by virtue of the principle of plenitude, includes a proposal to combine magnetic monopoles with tachyons [148]. It turns out, namely, that while the special theory of relativity does not explicitly predict the existence of neither (subluminal) monopoles nor tachyons, it explicitly predicts the existence of tachyonic monopoles. That is, if the special theory of relativity is rebuilt without assuming a priori that any material velocity is subluminal, then it can be shown that the electromagnetic field equations are fully symmetrical, with faster-than-light electrical particles appearing to behave as magnetic monopoles (see appendix II). "Tachyonic protons" and "tachyonic electrons" may then be identified with isolated south and north poles, respectively. Also these hybrid meta-particles have been loked for experimentally, but with the same, negative result.

The case of the monopole, or the tachyon, shows, furthermore, some similarity to another spectacular case in modern high-energy physics, viz. the discussion concerning the existence of the so-called quarks. These entities were predicted by Gell-Mann and Zweig in 1964 to be sub-nucleonic constituents carrying fractional charges. The early appeal of the quark model was related to its simplicity, tracing all hadronic matter down to three fundamental entities. Since then, the number of quarks - now endowed with different "colours" and "flavours" - has increased, leaving the model with a less aesthetical appeal. Physicists differ in their conception of what quarks are. Most physicists believe in quarks as a valid model, but there is no agreement as to whether they exist as physical objects or if they are merely mathematical constructions without a material existence. The experimental hunt for quarks shares the fate of the hunts for monopoles and tachyons. In 1969, an Australian group of physicists [121] reported that a few quark candidates had been observed among 6oooo cloud chamber tracks, but the discovery was immediately questioned and was not generally accepted. Despite numerous other experiments, there is today no experimental evidence for the physical existence of quarks [1o4].

The similarity between quarks and monopoles is not, perhaps, confined to this external level. It has been suggested that quarks should in fact be identified with pairs of magnetic monopoles [125]. By developing Dirac's old theory, it is argued that such an identification would explain the much discussed so-called "quark confinement", i.e. the fact that free quarks have never been observed but somehow seem to be confined within the domain of strong fields. A part of the recent interest in monopoles is not concerned with a development of Dirac's original idea, but rather with a variant worked out by Julian Schwinger. In this theory, so-called dyons, hypothetical particles which carry both magnetic and electric charge, are substituted for the Dirac poles. The charge quantization condition for dyons permits fractional magnetic and electric charges, thus leading to another unexpected contact with the quark

Monopoles, tachyons and quarks have been intensively sought for during the sixties and seventies, with particle accelerators, in the cosmic radiation, in stable matter and even in lunar soil. Despite the many experiments of an ever increasing sophistication and cost, and a few false discoveries, there is still no observational evidence for either monopoles, tachyons or free quarks.

The kind of physical theory, exemplified in Dirac's theory of the monopole, has a rather peculiar position from the viewpoint of theory of science: Dirac's monopoles can be verified, but never falsified. Any number of failed attempts to detect the Dirac pole will be insufficient to prove its non-existence. Dirac has never claimed that his monopoles are constituents of usual matter, or specified under which conditions they should be observable. He has only argued that they exist "somewhere in the world." Further, to satisfy Dirac's main argument for the monopole, that it implies quantization of the electric charge, the existence of only <u>one</u> monopole in our whole universe is sufficient. Due particularly to lack of knowledge concerning the rest mass of the monopole and the kind of interactions in which it participates, the theory is insufficiently complete, that a crucial experiment cannot be construed. Logically, Dirac's monopole theory is a purely s.c. existential statement, being of the same type as the classic example "there are black ravens".Such statements cannot be falsified and, adopting a falsificationist view, theories which essentially are existential statements should therefore not be considered as decent science. To Popper, strictly existential statements are non-empirical or "metaphysical", that is non-scientific([144], p. 69). Popperians, therefore, should reject Dirac's famous theory from methodological reasons alone. Unfortunately, monopoles have never been examined by philosophers of science.

The search for monopoles, tachyons and other elusive particles, ultimately derived from the principle of plenitude, may always be justified by their inevitable contribution to knowledge, irrespective of the outcome of the search. As argued by Bilaniuk and Sudarshan:

> "Regardless of the outcome of the search for tachyons, investigations in this field must invariably lead to a deeper understanding of physics. If tachyons exist, they ought to be found. If they do not exist, we ought to be able to say why not." ([5], p. 51).

In the light of the non-falsifiability of some of the modern theories of meta-particles, this argument seems to be invalid. Since there is no useful criterion to decide about the non-existence of e.g. monopoles, physicists may go on forever with their experiments without the theorists feeling obliged to give reasons to "why not". This situation seems, furthermore, to imply some consequences for the policy of pure research. Very considerable resources in manpower and finance have been allocated to the search for new particles, thought perhaps to exist somewhere in the universe. If these particles do not exist, or if they play hide-away with the physicists behind some distant galaxy, this will not prevent further experiments of a still more sophisticated and expensive sort, leading to neither experimental discovery or to increased "why not knowledge".

This ideology of pure science, including the hunt for new particles, is partly based on the usual science-for-the-sake-of-science attitude but also, as the pure science activities grow still more expensive and harder to defend politically by purely internal arguments, on hinted expectations of future applications. For this point the monopole story is most instructive. Having found his "monopole" track, Price did not fail to mention that "you might drive ships across the seas by putting a few monopoles in the ship and having the Earth's magnetic field tug it across the ocean" [17]. And in the press release accompanying the announcement of the discovery, it was even asserted that it could lead to such glorious applications as "new medical therapies in the fight against diseases such as cancer, and new sources of energy" (ibid).

In the afore-mentioned theory of monopoles with negative mass (see note 6), it was hinted that such knowledge might probably be useful to the construction of super-fast interstellar space rockets!

§ 7. DIRAC'S THEORY OF COSMOLOGICAL CONSTANTS.

In the mid-thirties, Dirac left for a time the main stream of quantum physics. His reasons probably were in the mathematical complexities of the current quantum field theory which, Dirac felt, obscured the basic physical ideas and made the whole field unattractive to him. Going his own ways, Dirac devoted in this period the major part of his intellectual resources to classical-relativistic electrodynamics (cf. §5) and cosmology, both fields being far away from the main stream physics of the period. His engagement in speculative cosmology was, in particular, a virtual roundabout. Until 1937, practically all of Dirac's work had concentrated on microphysics and general quantum theory.

Dirac's cosmological theory undoubtedly was inspired by current trends in the field, particularly those followed by Eddington and Milne. Eddington had for years tried to combine cosmology with quantum physics, the latter based on Dirac's relativistic theory. The wave equation of the electron was viewed as describing, not an isolated particle (which is a nonsensical notion in Eddington's philosophy) but the structural relation of the electron to the universe, being its comparison standard. On this idea, relativity cosmology was bridged with quantum theory, the study of galaxies merging with the study of electrons. In Eddington's programme, the constants of nature and their numerical values were of central significance. Taking e (the elementary charge), m (electron's mass), M (proton's mass), h (Planck's constant), c (velocity of light), G (graviation constant) and λ (the cosmical constant) to be the fundamental constants of nature, he attached a special significance to the dimensionless super-constants which may be constructed from these constants. These super-constants were in particular

$$\frac{M}{m}, \quad \frac{\hbar c}{e^2}, \quad \frac{e^2}{GMm} \quad \text{and} \quad \frac{c}{\hbar}\sqrt{\frac{Mm}{\lambda}}$$

Eddington wanted to deduce these constants from purely epistomological considerations and to connect them with the large number par excellence, the number of particles (nucleons) in the entire universe. Eddington found, for instance, the following relations:

$$\frac{e^2}{GMm} = \frac{\sqrt{3N}}{\pi} \quad \text{and} \quad \frac{\hbar c}{e^2} = N \cdot \frac{\lambda\hbar}{mc}$$

where N denotes the number of particles in the universe. The large numbers, N and the super-constants of the order of magnitude $\sqrt{N}$, were by Eddington considered to be true constants, independent of the cosmic expansion and the epoch, and demanding a rational understanding.

The approach initiated by Eddington in speculations concerning the significance of the cosmical constants, was followed also by other scientists. In the thirties, this peculiar kind of science fluorished. As a representative for the Eddington-inspired fashion we may mention Arthur Haas, who tried in a number of articles to deduce relationship between micro- and macrophysical constants of nature, thought to be of deep significance. He found, for example, that

$$\frac{e^2}{c} \cdot \frac{M}{m} = 2\sqrt{\frac{6}{5}} \cdot h \quad ,$$

and also ventured to derive the mass of the universe from basic assumptions [85]. Dirac was not the only quantum pioneer who felt the fascination of the Eddingtonian approach; In Germany, Jordan entered speculative cosmology in the same year, 1937 (see note 33).

Eddington's programme culminated, as earlier mentioned, in 1936 with the publication of <u>Relativity Theory of Electrons and Protons</u>. This book was preceded with one year by E.A.Milne's major work <u>Relativity, Gravitation and World Structure</u>, in which kindred problems were taken up in an equally non-orthodox way.[30] Milne was an Oxford astrophysicist of high standard and one of Dirac's former colleagues at Cambridge; he had been Dirac's supervisor for one term in 1925, when Fowler was on leave. Also Milne was concerned with the significance of the constants of nature, which he claimed were deducible from purely kinematical considerations, involving no appeal to empirical quantitative physics. But in contrast to Eddington, Milne's theory did not accord with general relativity and it also had no connection at all with quantum or atomic theory. Milne's cosmology was based on two assumptions, which were called the Cosmological Principle and the Dimensional Hypothesis. The first one states that the universe is everywhere uniform and isotropic about every point, so that the general physical phenomena are observed independently of the observer's placing in our universe. The Dimensional Hypothesis is a requirement that only dimensionless constants shall appear in cosmological theory. The main result, perhaps, of Milne's deductive system, was that the 'constant' of gravitation varies in proportion with the epoch:

$$G \sim t$$

Despite their differences, Eddington's and Milne's theories had in common their general approach to physics, first of all the modest power attributed to experimental reasoning as opposed to a priori arguments. Milne plainly rejected physics based on empirical and inductive methods as being irrational and inexact. Instead, he offered to "derive the laws of dynamics rationally .... without recourse to experience" ([132],p.1000). Such tunes had been whistled by Eddington for years. "The theory does not rest on .... observable tests," Eddington declared, "It is even more purely epistomological that macroscopic relativity theory. .... it should be possible to judge whether the mathematical treatment and solutions are correct, without turning up the answer in the book of nature. My task is to show that our theoretical resources are sufficient and our methods powerful enough to calculate the constants exactly - so that the observational test will be the same kind of perfunctory verification that we apply sometimes to theorems in geometry" ([65],p.3).

Dirac's entry in speculative cosmology in 1937 was, then, a natural participation in an already established tradition. The kind of problems Dirac took up were defined by Eddington and Milne. And their deductive approaches have probably appealed to Dirac who, in his work on quantun theory, favoured a related methodology. It was a distinct feature in Dirac's physics that he was never satisfied with accepting neither natural phenomena nor numerical constants as plain matters of fact, but he wanted to explain them from a deeper theory. The numerical value of the constants of nature had to be explained such as he had earlier strived to explain the quantization of the electric charge. This feature was congenial to Eddington's thinking, but Dirac thought that the explanation of the cosmological constants had to rest on different standards than those employed in atomic theory. Despite the resemblance, Dirac's cosmological theory was very different from both Eddington's and Milne's.

Dirac's point of departure was again the dimensionless constants, and particularly the very large ones. Of these, we may construct, among others, the following:

$$\gamma = \frac{e^2}{GMm} \simeq 2{,}3\cdot 10^{39}$$

$$\tau_0 = \frac{t_0}{e^2/mc^3} \simeq 7\cdot 10^{38}$$

$$\delta = \frac{c/H}{e^2/mc^2} \simeq 10^{39}$$

$$\mu = \frac{\rho(c/H)^3}{M} \simeq 10^{78}$$

In here, e,G,M,m and c have their usual meaning. t_0 is the age of our universe, measured in e.g. seconds. ρ denotes the mean density of matter in the universe and H is Hubble's constant, expressing the expansion of the universe by the spectral red-shift per unit distance. Dirac's cosmology was, as were most cosmological models in the thirties, based on a 'Big-Bang' assumption, such as worked out by Lemaître in 1931. The expressions given above may be assigned a physical meaning, even if this is mot really relevant to Dirac's arguments. Thus, γ is the ratio of the electric to the gravitational force between an electron and a proton; δ expresses the radius of the universe measured in terms of the classical radius of the electron; and μ is the mass of the universe in terms of the proton's mass, that is roughly the number of nucleons in the universe.

As to the numerical values, the constants τ_0, δ and μ were only very inaccurately known because of the rather rough estimates of t_0, ρ and H. In 1937 the best values for these quantities were $2 \cdot 10^9$y, $5 \cdot 10^{-31}$g$\cdot$cm^{-3}(for luminuous matter only) and $H = 2 \cdot 10^{-17}$sec^{-1}. The presently accepted values are $18 \cdot 10^9$y, $2 \cdot 10^{-29}$g$\cdot$cm^{-3} and $3 \cdot 10^{-18}$sec^{-1}. These numerical changes are not, however, damaging to Dirac's arguments, which are of a rather qualitative nature and not much sensitive to changes in experimental results. In this respect, Dirac's cosmological theory differed much from Eddington's, in which exact numerical agreement was highly favoured. For instance, Eddington deduced the number of particles in the universe to be exactly $2 \cdot 136 \cdot 2^{256}$ or approximately $3{,}15 \cdot 10^{79}$. By a curious, and of course purely accidental coincidence, the Eddington-Dirac figure for the total number of nucleons in the universe was essentially the one calculated by Archimedes 2200 years earlier![31]

The rough character of Dirac's reasoning may be illustrated by the constant τ_0 which denotes the age of the universe measured in 'atomic units', i.e. units of time fixed by the constants of atomic theory. From e, m, M, h and c several simple combinations with dimension of time may be constructed, with a maximum separation of circa 10^5:1. But which of these time units are chosen, does not really affect Dirac's numerological but non-quantitative reasoning. Dirac chose $\tau = e^2/mc^3$, which is the time required for light to traverse a classical electron's diameter. In this way he obtained $\tau_0 = t_0/\tau$ = circa 10^{39}.

For the mentioned dimensionless constants, Dirac observed that they are not scattered randomly as one might expect at first glance, but they exhibit a certain regularity in distribution: the numbers are distributed in three clusters, in the neighbourhood of 10^0, 10^{39} and 10^{78}, respectively. This gross regularity cannot be purely coincidental, Dirac claimed. "I now make the hypothesis that the clustering of the dimensionless numbers is a fundamental natural phenomena, which will hold for all time" ([132],p.1002). In his first, brief communication on the subject, Dirac called attention to the relationship between μ and τ_0, that is the relation $\tau_0^2 \simeq (10^{39})^2 = 10^{78} \simeq \mu$. This relationship was regarded, not as a curious coincidence but as an argument that the number of nucleons in the universe must increase with the square of the epoch:

$$N \sim \tau^2$$

That is, the hypothesis involves the assumption af a spontaneous creation of matter-energy. This assumption was, of course, difficult to defend on scientific grounds, as it lacked any theoretical or experimental support. Dirac, understandably, was not happy about this feature of non-conservation of matter, and shortly afterwards he decided to renounce it (see below). The principle of matter-conservation is, however, irreconcilable with the Large Number Hypothesis and in his post-war expositions of the theory, Dirac returned to his original view that matter must be continuously created in the universe. It should be remarked that the creation of matter, as required by the Large Number Hypothesis, har nothing to do with the continuous creation of matter as introduced by Hoyle and others in the Steady State theory.

Dirac was particularly fascinated by the approximate agreement between τ_0 and γ , an agreement which he presumed was due to "some deep connection in Nature between cosmology and atomic theory" ([36],p.201). By hypothesis γ was put equal to $k\tau$, k being of unit order of magnitude. This means that γ must vary proportionally with the time. Now the assumed time variation of γ may be due to $e = e(t)$, $G = G(t)$, $M = M(t)$ or $m = m(t)$, or some combination. Dirac rather arbitrarily took $G = G(t)$, probably because time variation of the atomic constants would imply radical changes in the fundamental atomic theory in which Dirac, the quantum physicist, kept so much confidence. So Dirac deduced that the gravitational 'constant' must decrease with the epoch according to

$$G \sim \tau^{-1}$$

The numerological speculations concerning G and N may resemble Eddington's, by whom, however, they were explained from epistomological principles. Dirac rejected Eddington's epistomologically based cosmology and removed the significance of the constants of nature from the realm of epistomology to the realm of history. The above deductions were considered as special examples of the

> "general principle that all large numbers of the order 10^{39}, 10^{78} ... turning up in general physical theory are, apart from simple numerical coefficients, just equal to t, t^2 ... where t is the present epoch expressed in atomic units." ([34], p.323)

In its most general form, Dirac formulated what he called the fundamental or Large Number Principle as follows: "Any two of the very large dimensionless numbers occuring in Nature are connected by a simple mathematical relation, in which the coefficients are of the order of magnitude unity" ([36],p.201). It should be remarked that Dirac did not, in contrast to Eddington, offer an explanation of the small dimensionsless constants such as the fine-structure constant.

In accordance with Eddington's spirit, Dirac expected a close connection to exist between the laws of the cosmos and the laws of the atom. Therefore he rejected Milne's dimensional hypothesis, which implies that the atomic constants shall not appear in cosmology, and replaced it with the Large Number Principle such as stated above. But Dirac accepted Milne's Cosmological Principle as a fruitful assumption. On this principle and on the Large Number Principle, Dirac considered Hubble's law. The distance between recessing galaxies varies with the epoch τ in some way, say $R = f(\tau)$, which may become a dimensionless number if expressed by an atomic unit of lenght such as e^2/mc^2. Hubble's constant may be expressed as $H = f'(\tau)/f(\tau)$ and was in 1937 given a value of circa $1{,}4\cdot 10^{-39}$. Since furthermore ρ was estimated to be of the same order for magnitude (if again expressed in atomic units), Dirac concluded from the Large Number Principle, applied to the reciprocal quantities, that

$$\rho = k\cdot H = k\cdot\frac{f'(\tau)}{f(\tau)}$$

By further assuming matter conservation we have that $\rho \sim f(\tau)^{-3}$, so that

$$f(\tau)^{-3} \sim \frac{f'(\tau)}{f(\tau)}$$

or, by integration

$$f(\tau) \sim \tau^{1/3}$$

which is the deduced law for the recession of the galaxies. The present age of the universe must be related to Hubble's constant according to

$$\tau_0 = \frac{1}{3}\,\frac{f(\tau)}{f'(\tau)} = \frac{1}{3}\cdot H^{-1}$$

Dirac admitted that this deduction did not agree well with the experiment data and might even appear inconsistent since τ_0 becomes smaller than the age of the earth. But he tried to circumvent this difficulty by pointing out that the age of the earth is calculated from data of radio-active decay and that, according to Dirac's ideas, the decay rate of atomic disintegration would also decrease with the epoch.[32]

Dirac's picture of the universe disagrees with the models which involve a finite universe, either from Steady-State assumptions or from the expansion assumption supplied with a contraction phase in the far future. In 1937, Eddington and most others stuck to finitism, such as was also implicitly assumed in Dirac's first note on the subject. But the assumption af a maximum size of the universe involves a large number, which does not vary with the epoch, and it thus contradicts the Large Number Hypothesis. According to Dirac's view, the universe goes on expanding, forever.

Even if Dirac worked out some further consequences of his theory, it was not founded on arguments of a more solid nature than the here mentioned aesthetic and numerological considerations. The heart of the theory was, and still is, the hypothesis that $G \sim \tau^{-1}$, and it was also this hypothesis that was particularly subjected to experimental testing in the following decades. Due to the enormous time span involved - Dirac's theory yields a half life of G of about 10^{10} years - this is a very difficult, but not a hopeless task. Surveying the various theories for fundamental constants and their experimental plausibility, Dyson concluded in 1972: "It is quite possible that all will fail [the coming experimental tests], and then it will be up to the speculative cosmologists, and up to Dirac in particular, to think of something new."([62],p.236)

Dirac's cosmological postulates, involving the time variation of G, is irreconcilable with Einstein's general theory of relativity which demands a true constant of gravitation. This was evident also in 1937, but Dirac did not then specify what the new field theory of gravitation, to replace Einstein's, should look like. It was only in 1973 that Dirac further developed his cosmology and explored its consequences for general relativity. In particular, Dirac attempted to connect his cosmological ideas with Weyl's old theory of a geometrically framed unification of electromagnetism and gravitation. This theory, which was generally rejected soon after its appearance, was, Dirac remarked "unrivalled by its simplicity and beauty."([54],p. 405). Dirac's extension of his cosmological theory was still based on his Large Number Hypothesis, in which Dirac expressed no less confidence that he did 35 years earlier. In 1973, Dirac justified his extended cosmological theory by arguing that it was "a further step in the direction of widening the group of transformations underlying physical laws". This echo from the 43 year old preface to the <u>Principles of Quantum Mechanics</u> was, in 1973 as in 1930, the ultimate heuristical principle for Dirac, constituting the highest form of mathematical beauty.

> "It appears as one of the fundamental principles of Nature that the equations expressing basic laws should be invariant under the widest possible group of transformations."([54],p. 418)

Dirac's 1937 contribution to cosmology caused more excitement in England than had his earlier, and scientifically much more important theories. It even reached the pages of <u>The Times</u>, which praised Dirac's theory, which "alters fundamentally our ideas of the structure of the universe and the nature of time"[21]. Many scientists and philosophers, however, objected strongly to Dirac's theory which was accused of perverting what was called the 'Galilean Method', i.e. the inductive method of building up theory to fit observations. Herbert Dingle, in particular, launched a front-attack against the unbalanced apriorist methods of the "Modern Aristotelians", to whom he counted Milne, Eddington and Dirac. Their approach, Dingle thundered, was nothing but a "combination of paralysis of the reason with intoxication of the fancy." "Instead of the induction of principles from phenomena we are given a pseudo-science of invertebrate cosmythology, and invited to commit suicide to avoid the need of dying"([21],p.786). Dirac's lining up with 'obscure' thinkers as Milne and Eddington clearly confused and annoyed Dingle: Dirac's theory was criticized "not as a source of infection but as an example of the bacteria that can flourish in the poisoned atmosphere; in a pure environment it would not have come to birth, and we should still have had the old, incomparable Dirac"([22],p.1012).

In the summer of 1937, Dingle's attack started a heated debate in Nature on the proper relationship between physics and philosophy. If Dirac was no longer "the old, incomparable Dirac", he was unaware of his sudden metamorphosis. In reply to Dingle, Dirac acknowledged that "the successful development of science requires a proper balance to be maintaimed between the method of building up from observations and the method of deducing by pure reasoning from speculative assumptions"([132],p.1001). He thought himself that he had kept this balance also in his cosmological theory, the lack of reliable observational material taken into regard. As usual, Dirac avoided to go into a philosophical discussion.

In reconsidering the debate, it is curious to note how both camps supported their arguments by reference to positivistic theories of knowledge. Dingle, and also the mathematician H. Jeffreys, accused Milne, Eddington and Dirac for neglecting evidences of experience and constructing decuctive systems without recource to experiments, thereby sinning against positivistic philosophy as taught by Mach and Pearson. Milne, for his part, justified his approach also in terms of Machian philosophy. He saw his cosmology as a result of "an extreme application of the principle of the economy of thought," indebted to "Einstein's principle of introducing only elements which can in principle be observed ([129],p.997). Indeed, if the debate showed nothing else, it did show how comprehensive a philosophical doctrine positivism is and how widely different approaches it may be used to legitimate.

Whatever the result of the debate, it may safely be said that Dingle's use of 'Aristotelianism' and 'Galileinism' as synonymous with apriorism and inductivism, respectively, was not very fortunate. As duly pointed out by G.J.Whitrow and G.D.Hicks, this has nothing to do with the actual history of science and philosophy. Neither Aristotle nor any genuine follower of Aristotle failed to recognize the importance of observation in scientific investigation, nor did they think that general principles can ever be established as deductive systems of the human mind. On the contrary, genuine Aristotelianism is rather associated with a common-sense empiricism and a pedestrian form of inductive reasoning. And for Galilei, Kopernicus, Kepler and other pioneers of modern science, they were, in fact, much more 'Aristotelian' than was Aristotle. The foundation of modern physics, such as created by Galilei, was much indebted to counter-inductive reasoning, and would then also fall under Dingle's criticism.

In regard to the undeniable speculative character of Dirac's cosmological theory and the very meagre experimental support for it, it may appear strange that it was welcomed rather favourably at its emergence and has continued to attract scientists' serious interest. There may be two reasons to this, I suggest, one based in the intellectual climate and one in the sociology of science.

First, Dirac's theory fitted nicely into the intellectual climate which predominated in England in the thirties. The philosophical positions in vogue in the late twenties and in the thirties had, for most of them, common tendencies of what may be described as objective idealism. Whitehead, Jeans, Russel and Eddington were, each in their own way, much occupied with the distinction between the world as it appears and the real world, which was considered as an entirely different character than the chaotic plurality of matter. Jeans' view, that "present-day science adds that .... much, and possibly all, that was not mental has disappeared, and nothing new has come in that is not mental"([103],p.298), characterizes well the spirit of the time. Also it is noteworthy that most of the influential British scientists' philosophies in the period gave a 'dematerialized' picture of the scientific enterprise, emphasizing the cultural and metaphysical aspects of the progress of physics, while neglecting its more pedestrian and traditional aspects, including technology. In this connection, it is tempting to ask if not this atmosphere in England was due to factors outside the scientific milieu; and more specifically, if not the situation may somehow be paralleled to the situation in Germany a decade earlier, such as it has been analyzed by Forman [79]. According to Werskey [164], a rather strong anti-science attitude was frequent in wide intellectual circles, picturing scientists as narrow-minded technicians, innately incapable of either creating or appreciating cultural, artistic or social values. The British environment in the thirties was no doubt less hostile to the mathematical sciences than was the environment of the Weimar Republic; The hostile attitude was less penetrating and the criticism weaker articulated, but it nevertheless constituted an important cultural trend. In the 1930s, science was ranked low in British culture. Without thorough studies it is impossible to say whether British scientists reacted to this relative hostility and isolation in a manner which bears any significant resemblance to the one described by Forman, i.e. adapted their science outlook to the hostile environment. It would be an interesting task for historians of science to investigate whether Forman's scheme is applicable, mutatis mutandis, to British science in the thirties. But irrespective of this, we may safely state that the soil in which Dirac's cosmological ideas grew had been well fertilized during the preceding years.

It is remarkable, that after the war, when the intellectual environment had changed rather drastically, Dirac's theory received no longer much attention. This change in interest was not rooted in any new data to weaken Dirac's idea, but rather seems to rest in a general change of scientific standards and values in the post-war era, whose atmosphere was no longer favourable to the 'New Aristotelians'. When Pascual Jordan[33] in 1952 developed der Diracsche Gedanke into a comprehensive and quantitative theory [109], he could state: "Soweit ich sehe, bin ich der einzige, der bereit gewesen ist, das von seinem Urheber selbst sogleich teilweise wieder aufgegebene Diracsche Weltmodell ernst zu nehmen und über seine genauere Präzisierung nachzudenken."(Ibid.;p.137)

Secondly, the persistent interest (or rather, the revived interest) in Dirac's theory - contrary to e.g. Milne's, which is largely forgotten today - is not only rooted in the increased technological possibilities for verification, but may also be ascribed to Dirac's outstanding position in the physical community. It is a well-known sociological fact that only highly merited scientists can afford to propound very unorthodox theories and still get them accepted as respectable science. When Dirac's cosmological speculations are today viewed as 'interesting', 'imaginative' and 'bold', this is partly due to the immense scientific reputation of their creator. If the very same theory had been suggested by some unknown physicist, not to say a layman, surely it would have been ridiculed or ignored and probably never been accepted for publication in any of the so-called recognized scientific periodicals.

If comparing the methods and principles of the cosmological theory with those appearing in Dirac's quantum theories, they appear most unlike each other. Much of this difference, however, is rooted in the different subjects, naturally inviting different approaches. In their essence, Dirac's general ideas about methods in physics are also manifest in his cosmological theory, although expressed in a different way. Dirac was not really a methodological convert, he just remained "the old, incomparable Dirac". Of course cosmology is, more than any other branch of science, the classical field for apriorism and speculations of all sorts; and these aspects were much more distinct in Dirac's cosmological theory than they were in his atomic theories. To illustrate the kind of reasoning which may be used in speculative cosmology à la Dirac, let me mention a more recent example. In 1961, the American astronomer Robert Dicke introduced a new cosmological theory which challenged Dirac's postulates. Commenting on Dicke's theory, Dirac argued: "On this [i.e. Dicke's] assumption habitable planets could exist only for a limited period of time. With my assumption they could exist indefinitely in the future and life need never end. There is no decisive argument for deciding between these assumptions. I prefer the one that allows the possibility of endless life."([220]; p.441)

To some extend, Dirac's cosmological theory shows a parallel to his theory of magnetic monopoles. In both cases he predicted, on purely theoretical grounds, phenomena to exist in nature. And in both cases his predictions were widely debated and subjected to experimental investigations which have not, until now, been able to supply any conclusive evidence pro or con. The general feeling among physicists is probably that Dirac's theories are both of them imaginative and maybe even fruitful, but nevertheless false.

Due to the kind of problem under investigation, the cosmological theory included virtually no mathematics and consequently Dirac's mathematical programme from 1930-31 was rather ineffective in this field. Dirac's cosmological theory does not start out from laws of nature whose mathematical consequences are deduced and sought interpreted physically. A pure principle of plenitude in the sense earlier employed, is therefore not the basis of the theory. But of course it rests heavily on principles and on a priori expectations of an aesthetic kind, such as do all cosmological theories. Although Dirac's reasoning does not involve a 'Leibnizian principle', it does involve a 'Pythagorean principle', viz. that numerical coincidences and regularities in nature are not merely coincidental but are manifestations of the interconnectedness and order of the laws in nature. Such numerological principles have a long and glorious tradition in the history of science. In his numerological reasoning, Dirac lined up with such thinkers as Pythagoras, Plato, Kepler, Mendeleev and Eddington.

In the authentic Pythagorean tradition the numerological principle was connected to whole numbers. Also Dirac imagined, in a true Pythagorean way, that the mysteries of the universe might ultimately be explained in terms of whole numbers. Two years after his entry in cosmology, Dirac speculated:

> "Might it not be that all present events correspond to properties of this large number[10^{39}],and, more generally, that the whole history of the universe corresponds to properties of the whole sequence of natural numbers? ... There is thus a possibility that the ancient dream of philosophers to connect all Nature with the properties of whole numbers will some day be realized." [38],p.129

To Dirac the Pythagorean principle was closely connected to his general mathematical philosophy, as he imagined a mathematical study of the whole numbers would probably lead to the new physical insight, required for an understanding of the cosmical mysteries. Dirac expounded his modern version of Pythagorean thinking in this way:

> "One hint for this development seems pretty obvious, namely, the study of whole numbers in modern mathematics is inextricably bound up with the theory of functions of a complex variable, which theory we have already seen has a good chance of forming the basis of the physics of the future. The working out of this idea would lead to a connection between atomic theory and cosmology." ([38],p.129)

Apart from the numerological, or Large Number principle, Dirac's reasoning may be said to involve also a 'historical principle'. The most characteristic feature in Dirac's theory was, as mentioned, that the constants of nature are not really true constants but in general they vary with the aging of our universe. This historicity of Dirac's theory is, however, a consequence of his numerological reasoning and has hardly been a primary aim for Dirac.

This is said because Lewis Feuer has made a point out of the fact, that the period in which Dirac's theory emerged was, he says, intellectually and socially marked by beliefs in social change and collapse of socialist society. This, Feuer argues, was the "socio-teleological principle" underlying Dirac's theory. The historical perspective, Feuer goes on, "carried the imprint of the Marxist notions that were agitating the universities in the thirties" ([75],p.393). However, Feuer applies his special version of socio-cultural externalism in a too uncritical way, which unfortunately destroys his otherwise interesting approach. It is true that classical Marxist thinking (Engels, in particular) emphasized that the laws of nature are historical phenomena, not eternal absolutes provided by God. And it is true that some British Marxists welcomed Dirac's theory as being 'dialectical' in this sense.[34] It may even be (although I much doubt so, cp. Werskey) that the British scientific community anno 1937 was moved by feelings of social change and therefore responded positively to Dirac's ideas. But the possible mechanism for acceptance of a theory by the scientific community is quite distinct from the (conscious or inconscious) motive which inspires the scientist to propose the theory, a distinction Feuer fails to recognize. Was social change really a part of Dirac's "emotional-ethical-aesthetic longings", which caused him, by some kind of thought association, to propound his cosmological model? If so, Feuer fails to provide any evidence at all for this suggestion. In fact, such a psycho-externalist explanation seems altogether unfounded in regard of Dirac's personality and social values. If anything, introversion and non-involvement in social and political affairs were characteristics of Dirac. He has never publicly expressed any interest at all in social matters, and he was hardly receptive to the Marxist agitation of those days. At any rate, Werskey, in his authoritative work on socialist tendencies in British academic science during the thirties, finds no reason even to mention Dirac's name [165]. Dirac's interest in science seems to have been totally unconnected with any interest in society. Being an ivory-tower scientist, Dirac was so contrary to e.g. Bernal, Hogben or Haldane as can be imagined. Furthermore, Dirac was in many ways a typical 'Cambridge Man', sharing the attitudes which Cantabrians were expected to uphold. The values generated from within the Cambridge culture included, according to Werskey, that

> "Political commitments were your own affair, as long as they did not impede your full participation in the activities of your chosen research community. (But politics were thought to be such an irrational enterprise that any overt preoccupation with them was bound to cast some doubt on your 'soundness'). Whatever else might be said about the life-style of Cambridge scientists, it was not one likely to inspire or sustain socialist convictions and practises" ([165],p.21)

So we may, in effect, safely discard Feuer's suggestion about a connection socialism-Dirac-cosmology.

§8. Physics and Mathematical Beauty, I.

Although there are distinct similarities between Eddington's philosophy of physics and Dirac's approach, and particularly so in the latter's cosmological theory, there are also great differences. First of all, Dirac was always a working physicist and he has never pretended to be a philosopher, nor has he ever been particularly interested in philosophical aspects of science. Eddington's deep concern with epistomological and ontological matters, his grandiose attempt to formulate a systematic Weltanschauung based on physical theory, was foreign to Dirac's mind. Dirac has never, so far as I know, been explicitly concerned with the concept of substance, the existence of the external world, the process of knowledge or other classical problems of philosophy, all of which played a central role in Eddington's thinking. While Eddington's theories departed from (or, were claimed to depart from) epistomological principles, Dirac has often stressed that this was not his approach. On the contrary, Dirac saw it as a gratifying feature for his approach, that arbitrary epistomological principles were replaced by sound mathematical reasoning. Mathematics, not philosophy, is the way forward, according to Dirac.

The only occasion at which Dirac really diverged from his favourite role as a pure theoretical physicist and engaged in discussions of a more direct philosophical nature, was in 1939, receiving the James Scott Prize. At this occasion, he delivered an address on "The Relation between Mathematics and Physics," [38] which contains the fullest account of Dirac's philosophy of physics. In 1939, Dirac had completed all of his greatest achievements in physics, and was still, only 37 years old, in his creative age. The message of the James Scott address may be considered as an outstanding physicist's reflections over the collected results of his science so far.

Why, Dirac asks, meets the mathematical-deductive method with such remarkable success in physics? This method, i.e. mathematically "to infer results about experiments that have not been performed," is obviously Dirac's own favourite method. The answer is, according to Dirac, that "this must be ascribed to some mathematical quality in Nature, a quality which the casual observer of Nature would not suspect, but which nevertheless plays an important role in Nature's scheme" (ibid.,p.122). That nature is endowed with a 'mathematical quality' is a familiar theme in the history of ideas, where such notions have been advocated by many Pythagorean minded philosophers and scientists since antiquity. We notice that Dirac apparently locates the mathematical quality in nature, that is, it is not merely attributed to the intellectual equipment of the scientist as a subjective quality, such as held by conventionalism.

Often the mathematical quality is identified with a principle of simplicity stating, for instance, that the laws of nature should be of a simple form. To Dirac, however, the connection between mathematics and physical knowledge goes far deeper than this. Although the principle of simplicity is a valuable instrument of research, modern science has demonstrated that it does not apply to natural phenomena in general. Newton's law of gravitation, for example, complies much better with the principle of simplicity than does the Einsteinian gravitation theory, which is only expressible in a complicated set of equations.[35] Still Einstein's theory is a better, deeper and more general theory, superior to Newton's. Mathematical beauty, not simplicity, is what characterizes the relativity theory and this is, according to Dirac, the key concept in the relationship between mathematics and physics. Therefore Dirac gives the theoretical physicists the advice:

> "The research worker, in his efforts to express the fundamental laws of Nature in mathematical form, should strive mainly for mathematical beauty. He should still take simplicity into consideration in a subordinate way to beauty ... It often happens that the requirements of simplicity and beauty are the same, but where they clash the latter must take precedence." (ibid.,p.124)

But what, one asks, *is* mathematical beauty? How does one recognize one theory to be 'beautiful', another to be 'ugly'? Dirac's answer is not very satisfying nor is it much informative: " [Mathematical beauty] is a quality which cannot be defined, but which people who study mathematics usually have no difficulty in appreciating" (ibid.,p.123). There can be no doubt, however, that Dirac associates mathematical beauty with generality and universality and also with what he calls "interesting groups of transformations". The Lorentz group, for instance, is much more beautiful than the Galilei group, because it is more general and includes the latter as a special case. But also less universal theories may possess great beauty. "Also non-relativistic quantum mechanics is a beautiful theory because it is complete".[36] Dirac's conception of mathematical beauty, on the other hand, does not necessarily include exactitude or mathematical rigour. On the contrary, Dirac has repeatedly stressed that exact equations and rigorous proofs should not be the physicist's prime concern, and that even approximations may contain a great deal of beauty. In this respect, he differed from some contemporary mathematicians with interest in physics, such as Weyl, Wiener and von Neumann, in whose conception of mathematical beauty, logic and rigour were emphasized.

Although Dirac, in the late thirties, tended to conceive the power of pure mathematics in an absolute, metaphysical way, he was too much of a physicist to let this tendency dominate his practical work. From his early experiences with engineering science he had learned to value theories based on approximations.

> "A problem like arranging the windings in the rotor of a dynamo involved some mathematics. It was a mathematics of whole numbers, but there was quite a bit of beauty in it ... I think that if I had not had this engineering training, I should not have had any success with the kind of work I did later on, because it was really necessary to get away from the point of view that one should deal only with results which could be deduced logically from known exact laws which one accepted, in which one had implicit faith." ([53],p.112)

This relaxed attitude to mathematical rigour was a general feature in Dirac's physics. Nowhere, perhaps, is it illustrated more lucidly than in Dirac's introduction of the so-called δ-function in 1926. An earlier version of this 'function' had been introduced and applied by another famous engineering-trained physicist, Oliver Heaviside, in 1893. Heaviside declared that "Mathematics is an experimental science, and definitions do not come first, but later on" ([101],p.227). Dirac may have agreed in this antipuristic attitude when he defined his δ-function by the unusual conditions that $\delta(x)=0$ for all $x\neq 0$ and $\int\delta(x)dx = 1$. He recognized that this does not define a proper function and that the procedure was far from satisfactory to mathematical purists. But, Dirac argued, "All the same one can use $\delta(x)$ as though it were a proper function for practically all the purposes of quantum mechanics without getting incorrect results" ([24],p.625). In the preface to his 1930 textbook, he used the opportunity to stress that mathematics is indeed a powerful tool but a tool it is, not an end:

> "Mathematics is the tool specially suited for dealing with abstract concepts of any kind, and there is no limit to its power in this field. ... All the same, the mathematics is only a tool ..." ([29],p.VI)

This antipuristic spirit, in general a profitable one to hold for physicists, has remained central to Dirac. In 1965, writing on quantum electrodynamics, he stated the case once again:

> "Even though we cannot aspire to complete rigor, we may set up a theory with a reasonable practical standard of logic, rather like the way engineers work. Engineers do not aim at complete rigor. In their calculations they continually neglect quantities which they believe can be neglected without invalidating their results, basing this belief on previous experience, or maybe just feeling. The physicist working with q-numbers will have to develop a similar feeling for what can be neglected." ([46],p.687)

To Dirac, the golden way of advance in theoretical physics is tightly connected with the development of pure mathematical theory, a view he justified by reference to the development of quantum theory, particularly the surprising emergence of non-commutative multiplication. The rules of mathematics are the free inventions of the mathematicians, while the laws of physics are those which Nature has chosen, Dirac maintained. We have previously (§4) mentioned that in some passages, Dirac's conception of science seems congenial to conventionalist views. This is evident from the preface to his 1930 textbook, and also the emphasis on mathematical beauty is a pet theme in Poincaré and other conventionalist authors. But it would be wrong to label Dirac as a conventionalist therefore. Considering the fundamental laws of physics to be objectively supplied by Nature, such as Dirac thought, is radically opposed to conventionalist thinking, according to which the laws of nature are purely man-made constructions.

Despite the different status of mathematical and physical theory, the development of modern physics has shown, Dirac argued, that there is a most perfect marriage between the rules which mathematicians find interesting on internal grounds, and the rules chosen by Nature. This provides the physicist with a "powerful new method of research", namely

> "... to begin by choosing that branch of mathematics which one thinks will form the basis of the new theory. One should be influenced very much in this choice by considerations of mathematical beauty ... Having decided on the branch of mathematics, one should proceed to develop it along suitable lines, at the same time looking for that way in which it appears to lend itself naturally to physical interpretation" ([38],p.125)

That is, purely mathematical considerations should lead the way and in particular one should pay attention to those branches of mathematics that have an "interesting group of transformations". This advice is, of course, based in the successes of relativity theory and quantum mechanics.

Exactly which branch of mathematics that is worthy to pursuit according to these very general criteria, is difficult to say. To be 'interesting' seems to be no less undefinable and subjective than 'beauty'. But Dirac recommends a close study of the theory of functions of a complex variable, a field which is of "exceptional beauty" and therefore likely to result in interesting physical insight. Dirac feels about projective geometry in the same way, while he does not consider fields as set theory or topology to be endowed with any particular beauty.

Dirac's emphasis on the power in physics of the theory of functions of a complex variable, such as he repeatedly stated in connection with his cosmological theory and also at later occasions, was probably influenced by his current work in quantum mechanics. A quantum dynamical system is usually represented by a function of real variables, whose domains are the eigenvalues of certain observables. In 1937, Dirac suggested to drop the condition of reality and to consider the variables as complex quantities, so that the representatives of dynamical variables could be worked out with the powerful mathematical machinery belonging to the theory of complex functions. Considering the dynamical variables as complex quantities, implies, of course, that they can no longer be associated with physical observables in the usual sense. Dirac admitted this loss of physical understanding, but did not regard the increased mathematical abstractness as disadvantageous: " ... we have, however, some beautiful mathematical features appearing instead, and we gain a considerable amount of mathematical power for the working out of particular examples" ([35],p.48). The use of the same mathematical theory in quantum physics as in cosmology, though applied in very different ways, was most appealing to Dirac's mind.

Following his mathematical philosophy, Dirac argued against any mechanistic scheme of physics. His objections to mechanicism were not rooted in the quantum mechanical indeterminacy of observation, but rather in a desire to keep the whole of nature inside the realm of mathematical treatment. "I find this position [mechanicism] very unsatisfactory philosophically, as it goes against all ideas of the Unity of Nature" ([38],p.126). Dirac's argument was: Classical mechanics operates with two types of parameters, a complete system of equations of motion and a complete set of initial conditions. With these provided, the development of any dynamical system is completely determined. However, while the equations of motion are amenable to mathematical treatment, the initial conditions are not. These are determinable only from observation. Hence an asymmetry arises: the description of the universe is separated in two spheres of which mathematical theory only applies to the one. To Dirac, this was an intolerable situation and contrary to the expectations of unity in nature. According to Dirac, all the initial conditions -the elementary particles, their masses and numbers, the constants of nature- must be subjected to mathematical theory. Dirac foresaw a mathematical physics of the future in which "the whole of the description of the universe has its mathematical counterpart" (ibid.,p.129). The phantom of classical mechanicism, Laplace's daimon, had to have recourse to the initial conditions in order to predict the development of the universe. In Dirac's philosophical vision, Laplace's daimon reappeared in an even more powerful version, being able to deduce <u>everything</u> in the universe by pure mathematical reasoning:

> "... we must suppose that a person with a complete knowledge of mathematics could deduce, not only astronomical data, but also all the historical events that take place in the world, even the most trivial ones. ... The scheme could not be subject to the principle of simplicity since it would have to be extremely complicated, but it may well be subject to the principle of mathematical beauty." (ibid.,p.129)

This mathematical credo of Dirac's may call in mind the ideas of James Jeans, rather than those of Eddington. Jeans shared with Eddington his general idealistic and conventionalistic tendencies, but his philosophical approach took a more rationalistic turn. In particular, Jeans' thinking was a virtual worshipping of mathematics, thought to be the first and last word in science as well as in the universe. "From the intrinsic evidence of his creation," Jeans wrote, "the Great Architect of the Universe now begins to appear as a pure mathematician" ([102],p.134). And furthermore, the product designed by the divine mathematician, <u>viz</u>. our universe, <u>is</u> itself a mathematical thought. Since the world is in essence mathematical, <u>all</u> phenomena can be described in mathematical terms. And if we are not able to do so, it is not because that there is anything unamenable to mathematical treatment but because our mathematical knowledge needs to be improved. "<u>The final truth</u> about a phenomenon resides in the mathematical description of it; so long as there is no imperfection in this our knowledge of the phenomenon is complete" (ibid.,p.141). That is, Jeans' mathematical philosophy included an epistomology as well as an ontology. The ultimate reality of the 'physical' world exists in the mathematical formulae by themselves. This last, ontological aspect was absent in Dirac's otherwise much Jeans-looking mathematical thinking.

§9. Physics and Mathematical Beauty, II.

Dirac's ideas about mathematical beauty and the directing role of pure mathematics in theoretical physics, such as formulated in 1939 and more preliminary in 1931, have been taken up also at later occasions. Apparently the development of physics since 1939 has only confirmed Dirac in his then stated philosophy. Among Dirac's many remarks on the subject, the following are particularly characteristic and illuminating :

> "I feel that a theory, if it is correct, will be a beautiful theory, because you want the principle of beauty when you are establishing fundamental laws. Since one is working from a mathematical basis, one is guided very largely by the requirement of mathematical beauty. If the equations of physics are not mathematically beautiful that denotes an imperfection, and it means that the theory is at fault and needs improvement. There are occasions when mathematical beauty should take priority over agreement with experiment. ... A beautiful theory has universality and power to predict, to interpret, to set up examples and to work with them. Once you have the fundamental laws, and you want to apply them, you don't need the principle of beauty any more ..." ([122],p.59)

> "It seems to be one of the fundamental features of nature that fundamental physical laws are described in terms of a mathematical theory of great beauty and power, needing quite a high standard of mathematics for one to understand it. ... One could perhaps describe the situation by saying that God is a mathematician of a very high order, and He used very advanced mathematics in constructing the Universe. Our feeble attempts at mathematics enable us to understand a bit of the universe, and as we proceed to develop higher and higher mathematics we can hope to understand the universe better. This view provides us with another way in which we can hope to make advances in our theories. Just by studying mathematics we can hope to make a guess at the kind of mathematics that will come into the physics of the future. ... It may well be that the next advance in physics will come along these lines: people first discovering the equations and then needing a few years of development in order to find the physical ideas behind the equations. My own belief is that this is a more likely line of progress than trying to guess at physical pictures." ([45],p.53)

> "One should keep the need for a sound mathematical basis dominating one's search for a new theory. Any physical or philosophical ideas that one has must be adjusted to fit the mathematics. Not the other way around. Too many physicists are inclined to start from some preconceived physical ideas and then to try to develop them and find a mathematical scheme that incorporates them. Such a line of attack is unlikely to lead to success. ... The reason I feel so strongly about the views I expressed above is because of the success I have had with them in the past. ... I learnt to distrust all physical concepts as the basis for a theory. Instead one should put one's trust in a mathematical scheme, even if the scheme does not appear at first

> sight to be connected with physics. One should concentrate on getting an interesting mathematics." ([57],p.1)

The content of these, and other remarks from Dirac's later years, agrees well with his philosophy of physics while being an active researcher. Evidently, Dirac is radically opposed to most traditional philosophies of science and especially so in the role attributed to mathematical theory versus physical reasoning and experimental work. The latter are, according to Dirac, of very little importance, if not detrimental to the creation of advanceful physics (so far as fundamental theory is concerned). That the laws of nature should be discovered without necessarily 'asking the nature', i.e. without performing experiments, and that experimental disagreements should be ignored in cases of theories possessing a formal beauty, this is quite heretical from the point of view of traditional theory of science. The emphasis on pure theory, and the corresponding disregard of empirical evidences is not, of course, Dirac's discovery. Considering God as a master-mathematician is thus a very old theme in the history of ideas, going back to Plato and of central importance to such different thinkers as Galilei, Kepler, Leibniz and Kant. In the thirties, like views were put forward by such notabilities as Weyl and Jeans.

Apart from his own works, Dirac was particularly inspired and impressed by the theory of relativity and Einstein's general approach to physics. Einstein's theory of relativity is the supreme paradigm of a beautiful and therefore correct physical theory. In the Einstein Centenary, Dirac praised the general theory of relativity in these tunes:

> "Let us now face the question, that a discrepancy has appeared, well confirmed and substantiated, between the theory and the observations. How should one react to it? How would Einstein himself have reacted to it? Should one then consider the theory to be basically wrong? I would say that the answer to the last question is emphatically No. Anyone who appreciates the fundamental harmony connecting the way nature runs and general mathematical principles must feel that a theory with the beauty and elegance of Einstein's theory <u>has</u> to be substantially correct. ... When Einstein was working on building up his theory of gravitation he was not trying to account for some results of observations. Far from it. His entire procedure was to search for a beautiful theory, a theory of

> a type that nature would choose. He was guided only by the requirement that his theory should have the beauty and elegance which one would expect to be provided by any fundamental description of nature." ([58],p.17)

In his appraisal of trans-empiricism and mathematical intuition, Dirac is far from alone. Many outstanding theoretical physicists have joined Dirac in considering Einstein's theory of gravitation as a theory which has to be true, because of its aesthetic and formal merits, and one which was created practically without the involvement of empirical reasoning. Among the quantum pioneers, Max Born was about the only one who held a widely different view.[37]

The Einsteinian way to physics and Einstein's general views on the relationship between experimental evidence and mathematical beauty, are undoubtedly in nice accordance with Dirac's philosophy of physics. To no less an extent than Dirac, Einstein repudiated positivistic and inductive-empirical ideals of science.[38] "There is no logical way to the discovery of these elementary laws. There is only the way of intuition," Einstein stated in 1918. And in 1946: "A theory can be tested by experience, but there is no way from experience to the setting up of a theory" (Quoted from [97],p.980). As a last example, in 1933 Einstein stated his current credo about experience versus mathematical thinking in the following way, being highly congenial to Dirac's view:

> "Nature is the realization of the simplest conceivable mathematical ideas. I am convinced that we can discover, by means of purely mathematical constructions, those concepts and those lawful connections between them which furnish the key to the understanding of natural phenomena. Experience may suggest the appropriate mathematical concepts, but they most certainly cannot be deduced from it. Experience remains, of course, the sole criterion of physical utility of a mathematical construction. But the creative principle resides in mathematics. In a certain sense, therefore, I hold it true that pure thought can grasp reality, as the ancients dreamed." ([67])

Also Dirac's suggestion that the theoretical physicist should not be overly concerned with experimental facts (or 'facts') in connection with a fundamental theory, and that beauty should dominate over experimental evidence, finds support in Einstein.

As is well known, Einstein had such *a priori* confidence in the formal power of his gravitation theory, that he didn't care much about the famous solar eclipse 'confirmation' of 1919: had the observations disagreed with the theory, Einstein should nevertheless have kept full confidence in the theory.[39] It is less well known, maybe, that Einstein took a similar position to what in 1906 seemed to be a significant experimental disproof of his special theory of relativity.[40]

What we shall call the Dirac-Einstein rule is, of course, strongly contrasting the more or less Baconian inspired methodological rules which have dominated science. According to these rules, we should avoid hypotheses to the greatest possible extent; in cases where a hypothesis clashes with an experimental fact, the fact should always be preferred. Scientists and philosophers have generally endorsed what has been called "Boyle's rule" ([1],p.128): we should never reject well established observation reports in favour of a hypothesis. But what does 'well established' mean? Even within the reign of Boyle's rule it would be unquestionable that in some cases observational evidence should be ignored for the sake of a hypothesis. On this willingness rests a substantial part of the success of modern, i.e. Galilean and post-Galilean, physics. It is well known that e.g. Copernicus' and Kepler's celestial theories were far from consistent with observational facts when they were created. That, however, did not bother these pioneers too much. They strongly felt their theories to be beautiful, and if they did not fit all the facts these astronomers were inclined to think that it was only too bad for the facts. Indeed, to Galilei the prime example of the victory of mathematical reasoning over the senses was the Copernican system: "I cannot sufficiently admire," Galilei declared, "the eminence of those men's wits, that have received and held it to be true ... as that they have been able to prefer that which their reason dictated to them, to that which sensible experiments represented most manifestly to the contrary ... I cannot find any bounds for my admiration, how that reason was able in Aristarchus and Copernicus, to commit such a rape on their senses, as in despite thereof to make herself mistress of their credulity."([81])

The Dirac-Einstein doctrine is, however, much stronger than the mere willingness to ignore certain immediate sense experiences for the sake of reason. It demands that not even extremely well corroborated experimental facts should be given priority over beautiful thories.

The doctrine that experimental truth sometimes should be sacrificed for mathematical and aesthetical considerations, was also shared by Hermann Weyl in his physical research. "My work always tried to unite the true with the beautiful," Weyl once said, "But when I had to choose one or the other, I usually chose the beautiful".[41] Like Dirac and Einstein, Weyl was attracted by the Platonic idea of a mathematical quality inherent in Nature's scheme, an idea which Weyl used in his attempt to reconcile Christian metaphysics with modern science. "The mathematical lawfulness of nature is the revelation of divine reason," Weyl proclaimed in 1932, "The world is not a chaos, but a cosmos harmonically ordered by inviolable mathematical laws" ([168], p.11 and p.21).

Weyl's purely mathematical approach to physics was distinct in most of his works on relativity and quantum mechanics. To mention just one example, Weyl proposed in 1929 a two-component relativistic wave equation for particles with zero mass and spin one half (see appendix I), considered to be a natural extension of Dirac's equation for the electron [166]. In 1929, no zero-mass particle was known (apart from the photon which has zero spin and obeys the electromagnetic equations), and consequently Weyl's equation appeared to be merely of mathematical interest. Indeed, Pauli rejected Weyl's equation as being "auf die physikalische Wirklichkeit nicht anwendbar" ([138],p.226). It was only much later that Weyl's equation, ignored for almost thirty years, was rehabilitated. In 1957, Lee and Yang, and Feynmann and Gell-Mann showed that the neutrino does, in fact, satisfy the Weyl equation. To regard the delayed success of Weyl's equation as a support for the Dirac-Einstein doctrine of sacrificing experimental truth for mathematical beauty, such as does Chandrasekhar [16], is, however, unjustified. When Pauli and other physicists rejected Weyl's theory, it was not on the ground that experimentally there was no particle available to fill out the potential mathematical being. It was because Weyl's equation is inconsistent with the principle of parity invariance, i.e. changes under space reflection. So, Weyl's equation was rejected exactly because it was not, in most phycisists' view, a beautiful equation, as it violated the aesthetically appealing principle of symmetry. This principle was introduced in quantum mechanics by Wigner in 1927 and quickly became a sacred principle in physics. Pauli, in particular, was always guided by a strong aesthetic belief in symmetry and conservation properties of the laws of nature. It was this belief which made him discard Weyl's equation. And it was the same belief which caused him to distrust the evidences in favour of parity nonconservation. As late as in 1957, after Lee and Yang had introduced parity nonconservation in weak interactions but just before Wu supplied the experimental proof, Pauli was convinced that the experiments would be in favour of symmetry. "I do not believe that the Lord is a weak left-hander," he wrote.[42] But Pauli was wrong. His absolute confidence in the aesthetic symmetry principle turned out, in this case, to have blocked his scientific imagination. With the new knowledge about the non-universality of parity invariance, the aesthetic standards changed; now Weyl's old equation was considered to be, after all, a beautiful equation.

Though not appearing in textbooks on methodology of science, the Dirac-Einstein rule is, in more or less strong variants, frequently practised by physicists working in fundamental theory. To mention only one example, Weinberg has reported on the reaction of Gell-Mann, when confronted with the experimentalists' claim to have disproved the theory of weak interactions. "I remember," Weinberg writes, "Murray Gell-Mann rising and suggesting to the meeting that since the experiments didn't agree with the theory, the experiments were probably wrong" ([161],p.41). In this case, it turned out, Gell-Mann did right in adopting the Dirac-Einstein attitude. The experiments were wrong.

Einstein's peculiar way of understanding fundamental physics, largely shared by Dirac, Weyl and others, has greatly attracted the interest of philosophers of science. Polanyi's view, that "the discovery of objective truth in science consists in the apprehension of a rationality which commands our respect and arouses our contemplative admiration," ([142],p.5) is in particular based on Einstein's theory of relativity. "Modern physics has demonstrated the power of the human mind to discover and exhibit a rationality which governs nature, before ever approaching the field of experience in which previously discovered mathematical harmonies were to be revealed as empirical facts," Polanyi states (ibid.,p.15). That fundamental physical discoveries are largely based on intuitive, aesthetic and trans-empirical reasoning, such as claimed by Polanyi, and that this view is in particular supported by Einstein's case, has been strenuously objected to by philosophers of a more positivistic orientation.[43] Also the claim has been firmly rejected by some socialist authors, who cannot find place for elements of beauty in their materialist conception of science.[44] Thus Bernal, in his widely read classic Science in History, takes pains to emphasize that relativity and quantum physics are no less the results of a materialist ontology and an inductive-empirical method than are other areas of physics. "Einstein's theories," Bernal writes, "were ... derived ultimately from experiments and gave rise to practical applications" ([4],p.744).

In the debate over trans-empirical and subjective factors' role in modern science, Dirac's discoveries and his philosophy of physics have not been subjected for consideration. This is unfortunate, because they would probably supply interesting material to the discussion. An examination of Dirac's approach to physics, such as the one outlined here, seems to add support to the view preferred by Polanyi.

Dirac claimed, as mentioned, that his doctrine of 'first mathematics, then physics' was not only supported by the development of relativity theory but also by the early development of quantum mechanics. Indeed, the doctrine was largely shared by other quantum pioneers, especially those who rejected Anschaulichkeit as a useful criterion in physical theory. The most outstanding representative of this view, Werner Heisenberg, wanted to base atomic physics on non-visualizable, mathematical models and to derive the physical insight from the mathematical formalism. Heisenberg was always fascinated by the simplicity and beauty appearing in the mathematical scheme of quantum mechanics. In discussions with Einstein in 1926, he admitted to be guided by an aesthetic criterion of truth, rooted in the mathematical scheme's simplicity, harmony and closedness ([93], p.75). But unlike some other believers in mathematical beauty, Heisenberg was convinced that the mathematically revealed beauty belongs to the objective nature itself, and is not merely an expression for the scientists' intellectual equipment.

Dirac felt that his conception of the methods and aims of physics was in close agreement with "Heisenberg's view about physical theory - that all it does is to provide a consistent means of calculating experimental results" ([40],p.18). Like Dirac, Heisenberg was methodologically indebted to Einstein, whose construction of the theory of relativity was seen as the paradigm for radical change in physical theory. In fabricating quantum mechanics, Heisenberg thought that he carried Einstein's research principles over in the atomic domain. From Einstein he learned that the mathematical structure of a physical theory leads to true knowledge of nature. In 1927, for example, Heisenberg stressed the methodological analogy between quantum mechanics and the general theory of relativity. Just as the new conceptions of space and time follow from the mathematics of the relativity theory, a radical change in the mechanical concepts, Heisenberg explained, "scheint aus den Grundgleichungen der Quantenmechanik unmittelbar zu folgen" ([91],p.173). Heisenberg was, however, not entirely unambiguous as regards the proper relationship between mathematics and physics, a fact that may be ascribed the opposite influences of Einstein and Bohr. The latter's qualitative and philosophical attitude to physics, so different from Dirac's, made a profound impact on the young Heisenberg. Bohr, Heisenberg has recalled, "feared that the formal mathematical structure would obscure the physical core of the problem, and in any case, he was convinced that a complete physical explanation should absolutely precede the mathematical formulation" ([92],p.98). This lesson, sharply contrasting the Dirac-Einstein thinking, was also significant in Heisenberg's rather eclectic approach to physics, side by side with his tendency to mathematical reasoning.[45]

For Dirac, he certainly recognized the depth of Bohr's thinking but his own temper and approach to physics was radically different from Bohr's Nur die Fülle führt zur Klarheit research programme. Dirac favoured simplicity and comprehensible equations and was mentally unable to appreciate Bohr's clarity-through-complexity philosophy.[46]

§ 10. THE CLAIM FOR BEAUTY IN SCIENCE.

Mathematical beauty, we have seen, was raised by Dirac to the status of a universal principle for research in theoretical physics. But wherein lies the source for beauty in science, as conceived by Dirac and others? Which criteria should be adopted for mathematical and physical beauty?

In the analysis of beauty, many ingredients have been proposed to constitute this non-elemental concept ([99],pp.81-88). Most often, harmony, realization of expectation, unsuspected relationships, completeness and simplicity are among these ingredients. "Unity in variety", was Coleridge's ultrashort definition of beauty, adopted by Bronowski to cover also beauty in science ([14],p.29). Chandrasekhar, the eminent astronomer, has proposed that exceptionality as well as conformity are both basic elements of scientific beauty. A theory is beautiful, says Chandrasekhar, if it is "exceptional to a degree that excites wonderment and surprise," and "scientific beauty is the proper conformity of the parts to one another and to the whole" ([16],p.29). Most scientists, I guess, would agree in such definitions. All the same, they are very inexact and may be interpreted very differently in specific cases.

Scientific beauty is different from mere simplicity, as Dirac pointed out. Also it should be distinguished from <u>elegance</u>, though it may be difficult to make a clearcut distinction. Dirac, in fact, did not care to do so (see the quotation, from 1954, below). Beautiful science often includes simplicity and elegance. Polanyi has distinguished between elegance and beauty in the following way: "We attribute an indeterminate range of veridical implications to a discovery possessing real beauty, but not to an innovation possessing mere elegance". ([143],p.105)

Returning to Dirac, his concept of mathematical beauty is largely equated with the no less vague concept 'interesting mathematics'. It implies, as mentioned, that beautiful theories must be universal and comprehensive. In Dirac's actual use of the principle of mathematical beauty it should, however, rather be understood as 'consistency with the fundamental principles of relativity and quantum mechanics' combined with some unstated principle of plenitude. It was in this form that Dirac tacitly applied mathematical beauty in 1928, 1930 and 1931. The claim that, in the words of Keats, "what the imagination seizes as beauty must be true - whether it existed before or not,"[47] is indeed a variant of the principle of plenitude: Not everything which can be rationally imagined exists truly in nature; but what is recognized as beautiful, does. Implicitly applying such a principle, Dirac construed his relativistic wave equation, predicted the existence of anti-electrons and argued for the existence of magnetic monopoles.

Whatever meaning is associated with Dirac's and others use of mathematical beauty, it remains a most airy and indeterminate concept. Being essentially an expression for psychological emotion, beauty, in science as well as in art, can only be grasped subjectively. Rational arguments in order to justify the quest for beauty, plenitude, simplicity etc. can hardly avoid to be, in the end, circular inferences or tautologies. Aesthetical principles are, by their very nature, outside the realm of rationalism. The fact that mathematical beauty is only recognizable to mathematical experts makes it, furthermore, a highly elitarian one (in which respect it differs from beauty in art). The elitarian aspects of the quest for beauty in science are particularly transparent in Poincaré, who considered the mathematical faculty, on which the apprehension of beauty in science depends, to be an innate quality, limited to a few great minds. Beautiful combinations of thought are, according to Poincaré, "those that can charm that special sensibility that all mathematicians know, but of which laymen are so ignorant that they are often tempted to smile at it"([141],p. 59). Poincaré's position has always been, and still is, popular among mathematicians who like to consider the mathematical sciences as arts, their primary justification being in the aesthetical pleasure obtained. As a modern exponent of Poincaré's view, let us quote Morris Kline: "The ultimate test of a work of art is its contribution to aesthetic pleasure or beauty. Fortunately or unfortunately, this is a subjective test and depends on the cultivation of a special taste. Hence the question of whether mathematics possesses beauty can be answered only by those who have studied the subject". ([112], p.523)

Though usually confined to theoretical work of a basic nature, scientific beauty may also be appreciated in purely experimental work. An outstanding example of the recognition of beauty in measurements may be found in Robert Millikan's research notebooks concerning his famous determination of the electronic charge, circa 1912. During the progress of his measurements, Millikan repeatedly expressed his pleasure with remarks such as "Beauty. Publish this surely, beautiful!"(see [96]). In Millikan's case, and in experimental work in general, beauty is associated with measurements which show agreement with the expected results and which are accurate and unambiguous. Beautiful experiments are equated with good, successful experiments. This is also why Michelson's celebrated aether-wind experiment was considered as a beautiful and crucial piece of experimental work only after the theory of relativity had been accepted. The obtained null-result was contrary to all expectations, a mystifying and disappointing result. Far from considering it as a beautiful experiment, Michelson called it a "failure".

In the advocacy of beauty in science it is sometimes overlooked that beauty is a concept which changes with the time and is subject to pressure from social and cultural changes. Feelings of beauty, simplicity and symmetry are always relative to the state of current knowledge and values. Some sense of beauty seems indeed to be a universal feature of human beings in all cultures and at all times; maybe it is a constituent feature in the structure of mind, an archetypical element in Jung's sense, such as argued by Huntley ([99],p.77). But this notwithstanding, the content of the concept of beauty varies widely in space and time. This is convincingly demonstrated by the history of art and ideas, and is no less recognizable in the history of science. Take such a concept as harmony, which seems to be closely related to beauty. Although not all kinds of beauty are endowed with harmony, most people would agree that harmonious theories or works of art are beautiful. But the alliance between harmony and beauty has not always been felt natural. During some periods, such as the early romantic age, disharmony, obscurity and even inconsistency were felt to be associated with beauty.

The classical example of beauty and simplicity in physical science, perhaps, is the ancient dogma of circular motion in astronomy. By aesthetical (but also religious and, at times, political) reasons this dogma was taken for granted from Plato to Tycho. After Kepler it lost its magic and was no longer associated with beauty in particular. This example may serve as a useful reminder of how aesthetical principles, when elevated to blanket canons, may inhibit conceptual innovations instead of advancing them. Very often, it turns out, beauty in science is in practise associated with 'consistency with current fundamental laws and standards', considered to be a minimum condition for beauty; that is, the element of conformity in Chandrasekhar's definition takes predominance. Therefore, a dogmatic use of the principle constitutes an element of conservatism and may prevent radical theoretic changes. If, say, Lorentz invariance is taken as an eternal absolute for beautiful physics, this jeopardizes to preclude any future break with the current relativistic paradigm. It is all too trivial to point out that if Einstein had stuck to Galilei invariance as a necessary part of beauty in mechanics, he would not have created the theory of relativity. In short, the principle of beauty must, as other aesthetical principles, be controlled by the symptoms of truth rather than be regarded as a factor of truth. Concerning the methodological principle of simplicity, Mario Bunge has said: "Ochkam's Razor -like all razors- must be handled with care to prevent beheading science in the attempt to shave off some of its pilosities. In science, as in the barber shop, better alive and bearded than dead and cleanly shaven" ([15],

p. 149). This warning may also hold for the principle of beauty.

The vagueness and subjectivity of aesthetical principles like beauty, seems to make them unfitted as methodological instruments of research. One might, of course, dream about some future science in which the principles of beauty and plenitude are transformed into scientifically precise formulations, in a way similar to the transformation of the principles of simplicity and sufficient reason into the later principles of least action. But this can only be speculations.

I think that Julian Schwinger came close to a reasonable judgment of the role of beauty in science, when he declared:

> "How beautiful it would be if the logically sound concepts of magnetic charge and dyons should prove to be at the heart of the subnuclear world! ... We have heard so much about the importance of beauty in physical theory. No doubt a correct theory will be beautiful (a cynic will say that our concept of beauty would evolve to make it so), but a merely beautiful theory has small chance of being correct. In short, beauty, as a criterion for validity, is necessary but not sufficient." ([123],p.426)

For my part, I believe that there is a good deal of truth in what Schwinger calls the cynical view. And, reasonable as Schwinger's point of view is, still it is curiously impotent as long as one cannot state exactly what distinguishes a beautiful theory from an ugly one.

In actual cases where scientists have to choose between hypotheses, the use of aesthetical criteria often turns out to be <u>de facto</u> identical to the use of more conventional standards, such as the degree of empirical reliability. It is largely with hindsight, then, that e.g. the principle of beauty turns out to be so fruitful. Maxwell's theory, for instance, has for a century been considered as a most beautiful theory; but when it emerged, the already existing theories -due to Ampére, Weber and Neumann- were considered as much more beautiful than Maxwell's 'horrible system'. This, at any rate, was the feeling on the Continent. It was only when it was realized that Maxwell's theory was a <u>better</u> theory, in the empirical sense, that the aesthetic status of the theory changed.[48] As a rule, new theories in physics have always been considered as endowed with less beauty than the old ones in periods of change. It appears, as pointed out by Kuhn ([114],p.154), that aesthetic factors are rarely decisive or of great importance in periods of revolutionary change. Usually it is only post factum -in the light of an already established paradigm- that new theories' aesthetic merits are recognized. That this is so, does not preclude, of course, that aesthetical factors play a decisive role in individual scientists' creation of ideas.

Consider the following statement, containing the essentials of Dirac's philosophy of physics:

> "With all the violent changes to which physical theory is subjected in modern times, there is just one rock which weathers every storm, to which one can always hold fast -the assumption that the fundamental laws of nature correspond to a beautiful mathematical theory. This means a theory based on simple mathematical concepts that fit together in an elegant way, so that one has pleasure in working with it. So when a theoretical physicist has found such a theory, people put great confidence in it. If a discrepancy should turn up between the predictions of such a theory and an experimental result, one's first reaction would be to suspect experimental error, and only after exhaustive experimental checks would one accept the view that the theory needs modification, which would mean that one must look for a theory with a still more beautiful mathematical basis". ([44],p.143)

In here, the operational difficulty of the principle of beauty is clearly, though unconsciously, present: "When a theoretical physicist" has found a theory which he finds beautiful, then "people put great confidence in it", we are assured. But this argument presupposes that the individual physicist's sense of mathematical beauty, his particular psychological constitution, is automatically shared by "people", i.e. the community of theoretical physicists. This assumption, however, is unwarranted. There is no evidence, whether from psychology or history of science, that physicists or mathematicians should have in common definite conceptions about the nature of beauty in their science.

When Dirac recommends theoretical physicists to start out with considering beautiful and interesting mathematics, this will be ineffective as long as there is no general consensus about which equations and mathematical techniques that are beautiful and interesting. And it is a plain matter of fact that beauty is understood differently by different scientists, even when they belong to the same scientific sub-culture and share scientific standards. Thus, most physicists regard group theory and topology as highly interesting and promising branches, while Dirac does not. Dirac, and most physicists with him, considered the idea of Minkowski spaces as particularly beautiful; Einstein, however, also being a believer in beauty in science, did not at all like Minkowski's ideas when they first appeared [146]. Other examples of conflicting views of the aesthetic quality associated with a theory, have been mentioned in the text: the theory of magnetic monopoles (§6) and Weyl's neutrino equation (§9) substantiate the point. Also we saw in §3 how Dirac, in deciding the nature of his 'holes', was induced by two aesthetic principles -the principle of unity and the principle of mathematical reasoning- which unfortunately pointed in different directions. And in 1936, when accepting Shankland's results largely because they supported his aesthetically based distrust in quantum field theory (§4), Dirac's sense of mathematical beauty betrayed him: Shankland's results soon proved to be wrong.

In effect, the principle of beauty and kindred aesthetical principles appear to be inapplicable as general guides for scientific research. One cannot build a policy of pure science on aesthetic criteria. All the same, the principle of beauty has, in the form of a priori expectations of what basic laws should look like, served as an indispensable instrument in the creation of many of our fundamental physical theories. Probably the principle of beauty has also led to numerous wrong theories, not recorded by the historians of science.

-oOo-

APPENDIX I : RELATIVITY QUANTUM MECHANICS

Although the really interesting things first happen when the electron is placed in an electromagnetic field, for the sake of simplicity we shall consider only the case of a freely moving particle.

In this case, the Schrödinger equation becomes

$$-\frac{\hbar^2}{2m}\,\Delta\psi = i\hbar\,\frac{\partial\psi}{\partial t} \tag{1}$$

with the corresponding eigenvalue equation

$$\Delta\psi + \frac{2m}{\hbar^2}E\,\psi = 0 \tag{2}$$

These equations may be conceived as the quantum mechanical translation of the classical energy expression $E = p^2/2m$ by means of the operator prescriptions

$$p_k \rightarrow \frac{\hbar}{i}\frac{\partial}{\partial x_k} \qquad \text{and} \qquad E \rightarrow i\hbar\frac{\partial}{\partial t} \tag{3}$$

The direct relativistic extension of this procedure is to depart from the relativistic energy expression

$$E^2 = \vec{p}^{\,2}c^2 + m_0^2c^4 \tag{4}$$

which leads to the Klein-Gordon equations

$$\Delta\psi - \frac{1}{c^2}\frac{\partial^2\psi}{\partial t^2} + \left(\frac{m_o c}{\hbar}\right)^2\psi = 0 \tag{5}$$

and

$$\Delta\psi + \frac{1}{\hbar^2c^2}\left(E^2 - m_o^2c^4\right)\psi = 0 \tag{6}$$

(5) is Lorentz invariant but is difficult to reconcile with the general interpretation of quantum mechanics which demands a first-order equation in the time derivative. In the case of a Coulomb field present (6) gives a wrong fine-structure for the hydrogen spectrum.

The alternative procedure is to add correction terms of the order v^4/c^4 to the Hamiltonian in (2). This accounts for first-order relativistic effects, and also for spin effects, but leaves the theory in a non-invariant form.

A theory which is consistent with (a) the principle of relativity, and (b) the principles of general quantum mechanics, requires to be based on an equation which is (a) Lorentz invariant in x_k and t, or in p_k and E, and (b) of the first order in the time derivative (the energy).
This suggests, was Dirac's argument, to write (4) as

$$\frac{E}{c}\,\psi = \sqrt{(m_o c)^2 + p_1^2 + p_2^2 + p_3^2}\;\psi \qquad (7)$$

where E and p_k are still operators, given by (3). In order to conform with requirement (b), the square root has to be linearized:

$$\sqrt{p_1^2 + p_2^2 + p_3^2 + (m_o c)^2} = \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \beta(m_o c) \qquad (8)$$

This is not possible if the α, β coefficients are usual numbers but may be carried through if the coefficients are taken to be 4x4 matrices satisfying the relations

$$\begin{aligned} &\alpha_i\alpha_k + \alpha_k\alpha_i = 2\delta_{ik} \\ &\alpha_i\beta + \beta\alpha_i = 0\,, \quad \alpha_i^2 = \beta^2 = 1 \end{aligned} \qquad (9)$$

These so-called Dirac matrices may have the explicit form

$$\vec{\alpha} = \begin{pmatrix} \mathbf{0} & \vec{\sigma} \\ \vec{\sigma} & \mathbf{0} \end{pmatrix} \qquad \beta = \begin{pmatrix} \mathbf{1} & \mathbf{0} \\ \mathbf{0} & \mathbf{1} \end{pmatrix} \qquad (10)$$

where $\vec{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$ are the Pauli spin matrices and $\mathbf{1}$ is the 2x2 unit matrix. (7) may now be written

$$(p_o + \vec{\alpha}\cdot\vec{p} + \beta m_o c)\psi = 0 \qquad (11)$$

with $p_o = -E/c$. (11) is Dirac's equation for an electron. Its relativistic invariance is not evident in (11), but can be proved. Since the coefficients define 4x4 matrices, the wave function must have four components:

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}$$

The simplest case of Dirac's equation is the one of an electron at rest:

$$(p_o + \beta m_o c)\psi = 0$$

which, if using the β from (10), is immediately separable in the four equations

$$(p_o + m_o c)\psi_1 = 0, \qquad (p_o + m_o c)\psi_2 = 0$$

$$(p_o - m_o c)\psi_3 = 0, \qquad (p_o - m_o c)\psi_4 = 0$$

with solutions

$$\psi = \exp\left(\frac{m_o c^2}{i\hbar}\cdot t\right) \qquad \text{for} \qquad \psi_1 , \psi_2$$

$$\psi = \exp\left(-\frac{m_o c^2}{i\hbar}\cdot t\right) \qquad \text{for} \qquad \psi_3 , \psi_4$$

Of these solutions, the first two correspond to positive energy, the latter two to negative energy.

For electrons in motion, similar results appear. There are four linear independent solutions of the plane wave form

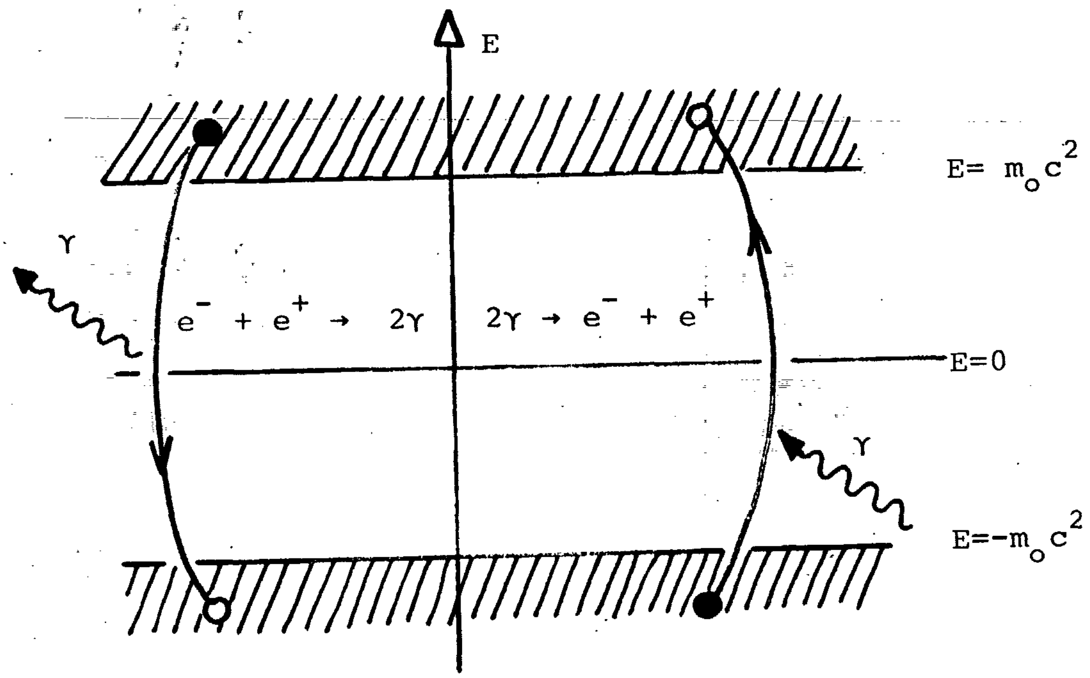
$$\psi_\mu = a_\mu(\vec{p})\exp\{ \frac{i}{h}(\vec{p}\cdot\vec{x} - E\cdot t)\} , \quad \mu = 1,2,3,4$$

The eigenvalue equations show that there are non-vanishing solutions only for

$$E = \pm c \sqrt{m_o^2 c^2 + \vec{p}^2}$$

such as would be expected also from the classical expression (4). For $\mu = 1$ and 2 the positive sign holds, while $\mu = 3$ and 4 are associated with negative energies. It can further be shown that the wave system which represents the electron, is complete only if the negative-energy solutions are included.

The energy spectrum consists of two allowed continuous areas, separated by the energy $2m_o c^2$. Annihilation and pair creation in



Dirac's picture are shown in the figure. For pair creation, the situation is that a negative-energy (negatively charged) electron in the 'sea' E<0 absorbs a photon of $E>2m_0c^2$ and is transferred to a E>0 state. As a result, a hole is left in the sea. The observable effects of this transition are

$$E_{obs} = E - E_{vacuum} \qquad \text{and} \qquad Q_{obs} = Q - Q_{vacuum}$$

where E and Q are the energy and the charge of the Dirac sea. Then we have

$$E_{obs} = (E_{vacuum} - (-|E|)) - E_{vacuum} = |E|$$

and

$$Q_{obs} = (Q_{vacuum} - (-|e|)) - Q_{vacuum} = |e|$$

thus, the appearance of a hole in the Dirac sea looks like the creation of a positive-energy particle with a positive elementary charge, a positron.

Another representation, also satisfying the relations (9), is

$$\vec{\alpha} = \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & -\vec{\sigma} \end{pmatrix} \qquad \beta = \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix}$$

using this representation, and writing

$$\psi = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}$$

where ϕ_1 and ϕ_2 are now two-component wave functions, (11) may be written as two equations, only coupled through the mass term:

$$\begin{aligned} (p_o + \vec{\sigma}\cdot\vec{p})\phi_1 - m_o c\phi_2 &= 0 \\ (p_o - \vec{\sigma}\cdot\vec{p})\phi_2 - m_o c\phi_1 &= 0 \end{aligned} \qquad (12)$$

To describe a mass-less spin ½ particle, only a two-component wave function is necessary, as (12) are now decoupled. For $m_o=0$ one gets

$$\begin{aligned} (p_o + \vec{\sigma}\cdot\vec{p})\phi_1 &= 0 \\ (p_o - \vec{\sigma}\cdot\vec{p})\phi_2 &= 0 \end{aligned}$$

It was the latter of these equations, that is

$$\frac{1}{c}\frac{\partial\phi}{\partial t} = \vec{\sigma}\cdot\nabla\phi$$

which was considered by Weyl in 1929. Weyl's equation is relativistically invariant and its wave function satisfies a continuity equation. It is not, however, invariant under space inversion, such as is the Dirac equation.

APPENDIX II: MONOPOLES IN ELECTRODYNAMICS

Maxwell's field equations in vacuum (CGS units) are

$$\nabla\cdot\vec{E} = 0\ , \qquad \nabla\times\vec{B} = \frac{1}{c}\frac{\partial\vec{E}}{\partial t} \tag{1.a}$$

$$\nabla\cdot\vec{B} = 0\ , \qquad \nabla\times\vec{E} = -\frac{1}{c}\frac{\partial\vec{B}}{\partial t} \tag{1.b}$$

That is, the system of equations are completely symmetric under the interchanging of electricity and magnetism, viz.

$$\vec{E} \rightarrow \vec{B} \qquad \text{and} \qquad \vec{B} \rightarrow -\vec{E}$$

In the presence of charges the symmetry is no longer complete. The Maxwell equations are now

$$\nabla\cdot\vec{E} = 4\pi\rho\ , \qquad \nabla\times\vec{B} = \frac{4\pi}{c}\vec{j} + \frac{1}{c}\frac{\partial\vec{E}}{\partial t} \tag{2.a}$$

$$\nabla\cdot\vec{B} = 0\ , \qquad \nabla\times\vec{E} = -\frac{1}{c}\frac{\partial\vec{B}}{\partial t} \tag{2.b}$$

In (2) there are no magnetic sources, hence the asymmetry. If magnetic monopoles are assumed by hypothesis, (2.b) takes the form

$$\nabla\cdot\vec{B} = 4\pi\rho'\ , \qquad \nabla\times\vec{E} = -\frac{4\pi}{c}\vec{j}' - \frac{1}{c}\frac{\partial\vec{B}}{\partial t} \tag{3}$$

where ρ' is now the magnetic charge, $\vec{j}'$ its corresponding current density. Now symmetry is established.
It is a pleasant feature of the Maxwell equations (2) that they may be reproduced by means of electric and magnetic potentials, ϕ and $\vec{A}$, defined as

$$\vec{B} = \nabla\times\vec{A} \qquad \text{and} \qquad \vec{E} = -\frac{1}{c}\frac{\partial\vec{A}}{\partial t} - \nabla\phi \tag{4}$$

This is not possible, however, in the presence of monopoles. Since

$$\nabla\cdot(\nabla\times\vec{A}) = 0$$

holds identically, (3) can never be reconciled with (4). I.e. monopole electrodynamics cannot be stated in terms of the electromagnetic potentials alone.

Hamiltonian and Lagrangian formulations of electrodynamics depend crucially on the existence of the electromagnetic potentials (4), not on the field quantities themselves. For example, the classical Lagrangian and Hamiltonian for a particle with charge q in an electromagnetic field are given by

$$L = \tfrac{1}{2}mv^2 + q\vec{v}\cdot\vec{A} - q\phi$$

and

$$H = \frac{1}{2m}(\vec{p} - q\vec{A})^2 + q\phi$$

Now the conventional way of transferring electrodynamics to quantum mechanics goes via a Hamiltonian or Lagrangian formalism. Since these formalisms are based on the electromagnetic potentials, quantum electrodynamics cannot contain monopoles as long as (4) is maintained.

If only electrified particles are allowed, but if in addition to the usual subluminal charges $(\rho,\vec{j})$ also superluminal charges $(\rho^*,\vec{j}^*)$ are introduced, then the field equations take the form:

$$\nabla\cdot\vec{E} = 4\pi\rho \quad , \qquad \nabla\times\vec{B} = \frac{4\pi}{c}\vec{j} + \frac{1}{c}\frac{\partial\vec{E}}{\partial t}$$

$$\nabla\cdot\vec{B} = -4\pi\rho^* \quad , \qquad \nabla\times\vec{E} = \frac{4\pi}{c}\vec{j}^* - \frac{1}{c}\frac{\partial\vec{B}}{\partial t}$$

That is, completely symmetric equations with the superluminal electric charges appearing as the magnetic monopoles in (3).

APPENDIX III : MONOPOLES IN QUANTUM MECHANICS. DIRAC'S REASONING.

Consider the wave function for, say, en electron: $\psi = \psi(x_j, t)$. ψ is only determined within multiplication of an arbitrary phase factor $e^{i\gamma}$; for ψ and

$$\Psi = e^{i\gamma}\psi \tag{1}$$

give the same probability distribution:

$$|\psi|^2 = |\Psi|^2 \tag{2}$$

In general the phase γ will be a function of position and time. We may take γ to be a non-integrable function. In this case, γ does not have a definite value at each point, but it has a definite change in value from one point to a neighbouring point. That is, it has definite derivatives

$$\frac{\partial \gamma}{\partial x_j} = \kappa_j \qquad \text{or} \qquad \nabla\gamma = \vec{\kappa}$$

For non-integrable phases, the change in phase around a closed loop

$$\Delta\gamma = \oint \kappa_j dx_j$$

will in general be different from zero.
By Stoke's theorem we get

$$\Delta\gamma = \int_S (\nabla \times \vec{\kappa})_n dS$$

where S is a surface whose boundary is the loop considered. The wave equation, whether relativistic or non-relativistic, involves the momentum operator $p_j = -i\hbar\partial/\partial x_j$. From (1) we get

$$-i\hbar\frac{\partial \Psi}{\partial x_j} = e^{i\gamma}(-i\hbar\frac{\partial}{\partial x_j} + \hbar\kappa_j)\psi = (-i\hbar\frac{\partial}{\partial x_j} + \hbar\kappa_j)\Psi \tag{3}$$

This means that ψ and Ψ will not, despite (2), satisfy the same wave equation. If ψ satisfies any wave equation involving

p_j, then Ψ will satisfy the corresponding equation in which p_j has been replaced by $p_j + \hbar\kappa_j$. This situation resembles what happens if we introduce an electromagnetic field. In this case, the equation of motion becomes the same as in the case of no field if only we make the substitution

$$p_j = -i\hbar\frac{\partial}{\partial x_j} \rightarrow -i\hbar\frac{\partial}{\partial x_j} + \frac{e}{c}\cdot A_j$$

where A_j is the vector potential and e is negative for the electron. Compared with (3), this means that the introduction of the non-integrable phase factor amounts to the same effect as introducing a magnetic field for which

$$\frac{e}{c}\cdot A_j = \hbar\kappa_j$$

We then have

$$\Delta\gamma = \frac{e}{\hbar c}\int_S (\nabla\times\vec{A})_n dS = \frac{e}{\hbar c}\int_S B_n dS \tag{4}$$

I.e., the magnetic flux going through the loop is connected to the change in the phase γ when going round a loop.
Next consider the change

$$\gamma \rightarrow \gamma + n\cdot 2\pi$$

This change leaves the wave function completely unaffected. But it does affect the result (4) since $\Delta\gamma$ will be different for different values of n, while the flux is completely definite. This demands a generalization of (4), namely

$$\Delta\gamma + n\cdot 2\pi = \frac{e}{\hbar c}\int_S B_n dS \tag{5}$$

where n is some definite, but unknown, integer.
Usually n is zero in (5) for very small loops: the magnetic flux will be close to zero and the change in the continuous wave function's phase must also be very small. Then we return to (4). But this argument breaks down in the case of vanishing wave functions, or for regions in space where ψ vanishes. In the case of $\psi=0$, for instance, γ is completely undetermined; for ψ close to zero, even small changes in ψ may correspond to

appreciable changes in γ, so that n has to be non-zero in (5) while the flux is still close to zero. Since ψ is a complex function, $\psi = \psi_1 + i\psi_2$, its vanishing will require two conditions, one for ψ_1 and one for ψ_2. In general the points at which ψ vanishes will therefore lie along a line, called a nodal line. For small loops around a nodal line $\Delta\gamma$ will then be equal to $2\pi n$, with n undetermined but non-zero.

A large loop may be treated by dividing it up into small loops lying in a surface whose boundary is the large loop. The flux passing through the large loop will equal $\Sigma\Delta\gamma$ for the small loops plus a contribution of $\Sigma 2\pi n$ from each nodal line cutting the surface. That is,

$$\Delta\gamma = 2\pi\Sigma n + \frac{e}{\hbar c}\int_S B_n dS$$

where the summation is over all the nodal lines, one term for each line. For a closed surface $\Delta\gamma$ must vanish, because the boundary perimeter shrinks to zero. Then

$$2\pi\Sigma n = \frac{e}{\hbar c}\int_S B_n dS$$

If one or some of the nodal lines have their end points inside the closed surface, Σn will not vanish and there will be a net magnetic flux crossing the surface

$$\int_S B_n dS = \frac{2\pi\hbar c}{e} n \qquad (6)$$

This magnetic flux implies the existence of a source, a magnetic monopole. Comparing with Gauss' law in electrostatics

$$\int_S E_n dS = 4\pi q$$

(CGS units) the strength of the monopole must be

$$\mu = \frac{\hbar c}{2e} n$$

The referred argument, essentially Dirac's 1931 argument, rests on the use of the magnetic potential defined by $\nabla\times\vec{A}=\vec{B}$. Since this is inconsistent with (6), electrodynamics requires some modification. The equation $\nabla\times\vec{A}=\vec{B}$ may be supposed to fail at just one point on the surface, where it is cut by the nodal line. This line of points, extending outward from the pole, is the so-called Dirac string. Dirac proposed to modify electrodynamics so that a term is added to $\nabla\times\vec{A}$; this term will vanish everywhere except in regions where the string passes.

-oOo-

NOTES.

1. No full biography of Dirac has as yet appeared. Mehra [122] gives much valuable information on Dirac's life and scientific work until the beginning of the thirties. Also Dirac's autobiographical sketches, especially [53], are valuable in this connection. Both of these works are confined, however, to Dirac's early work in quantum theory and do not go beyond 1933.

2. In a letter to Wilhelm Wien of 25 August 1926. Cited in translation from [79],p.104.

3. Quoted from [122],p.52. This is only one of the numerous anecdotes about Dirac, all of them expressing the peculiar unsocial, introvert and one-sidedly physical-mathematical character of his personality. For a collection of Dirac anecdotes, see [123], esp. pp.805-819.

4. "Ode to a Grecian Urn", written 1819. These four innocent lines, finishing Keats' poem, have caused a diversity of interpretations among literary critics who have disagreed as to the meaning and significance of Keats' "Beauty is truth, truth beauty" phrase. See[118].

5. For this problem, Dirac's way to the solution and the subsequent discovery of the positron, Russell Hanson has provided a very interesting account. See [88], ch.IX.

6. This is the general view, at any rate. It has not, however, prevented some physicists to speculate about the physical reality of negative-energy particles. They appear, for instance, in connection with tachyon theory (cf. §6) and were discussed by Sommerfeld in 1905. More recently, at least one author has suggested that magnetic monopoles are in fact negative-mass, and then negative-energy particles. Not in Dirac's imaginary sense, but particles endowed with a real physical existence in nature. See [171].

7. For this reason, Gamow dubbed them 'donkey-electrons', referring to the strange behaviour that the harder they are pushed, the slower they will go! If the paradoxical behaviour is not, in itself, sufficient to exile negative-energy particles from physics (cf. that tachyons are supposed to behave in a similar way), it may further be shown that charged negative-energy particles lead to direct inconsistency. Cf. Dirac's argument in 1930: "Although a negative-energy electron moves in an external field as though it has a positive charge, yet, as one can easily see from a consideration of conservation of momentum, the field it produces must correspond to its having a negative charge, e.g., the negative-energy electron will repel an ordinary positive-energy electron although it is itself attracted by the positive-energy electron" ([27], p.362).

8. In letters from Heisenberg to Jordan, 22 January 1929, and from Pauli to Klein, 18 February 1929. Quoted from [149], p.82.

9. 'Physically existing' is here to be understood in the unsophisticated sense adopted by every practising scientist, viz. that such entities must be capable of detection by our usual instruments, that is, exchange energy with the ordinary particles which compose our measuring devices. For scientists' purpose, entities with which we cannot communicate by experimental interaction -such as the electrons with negative energy in Dirac's theory- do not exist. Philosophically speaking, this is not, of course, a very satisfactory answer to the difficult problem about physical existence. Also, the standards among physicists for accepting entities as physical particles have changed rather radically during the latter fifty years, becoming less severe and accepting still more indirect evidences. See [61].

10. Gaston Bachelard, in particular, has drawn philosophical implications from Dirac's hole theory, claimed to support his ideas of a "philosophy of no" and being an example of "dialectical sur-rationalism"; see [3],p.59ff. In Bachelard's version, Dirac's theory is "idealistic" and "de-realized from its very start". Bachelard makes what I feel is an unwarranted point out of what is said to be the negative masses and negative energies appearing in Dirac's theory. The negative mass, Bachelard states, is a "dialectical" version of the positive mass, an entity which was inconceivable in pre-quantum physics and one which was only raised to legitimacy with the discovery of the positron. Apart from the objections that may be raised against Bachelard's poetical mystification of Dirac's theory, it is not true that negative mass was inconceivable in earlier "materialist" physics. Negative mass was, in fact, considered at one stage of the phlogiston theory, and it also appeared in connection with the discussions of superluminal electrons around the turn of the century.

11. For some reactions to the hole theory, see [88].

12. The text has "γ-Strahl-Protonen", which obviously is a misprint.

13. See[11],p.316. The letter dates from 9 December, 1929.

14. In his memoirs, Dirac has repeatedly stated that in 1930 "everyone felt pretty sure that the electrons and the protons were the only elementary particles in Nature" ([53], p.145). To be sure, Dirac admits, there was in 1930 the idea of the neutron but "people did not really have much faith in the existence of neutrons. It seemed to everyone self-evident that as there were two kinds of electricity, there should be just two kinds of particles to carry them" (ibid.). In this case, as in others, Dirac's

memoirs do not correspond exactly to the actual state of affairs. It is not quite true, then, that physicists in 1930 stuck tenaciously to the two-particle view. Apart from the well-known photon, the hypothetical neutron was in many quarters, and not least in Cambridge, considered to be a real elementary particle, only waiting for its experimental discovery. Rutherford, for sure, believed firmly that neutrons were not merely particles of the fantasy. See [72]. That the claim of a general belief in the two-particle view anno 1930 is untenable, is also supported by the retrospective comments by Oppenheimer, Bohr and Mott, quoted in [88],p.223.

15. For the history of particle creation and annihilation, see [13].

16. "Es ist naheliegend, zu erwarten, dass von den beiden Komponentenpaaren der Diracschen Grösse das eine dem Elektron, das andere dem Proton zugehört" ([166],p.332). Dirac acknowledged this suggestion of Weyl in 1930.

17. This is not quite true, for in Pauli's first announcement of his idea, appearing in an open letter of 4 December 1930, he thought about the neutrino as a light constituent of the nucleus, a neutral electron, and not as a massless particle, a spin ½ photon. See [149], p. 87. Pauli soon recognized that the neutrino could not reside within the nucleus but he continued to think about it as endowed with a mass, although very small.

18. [137],p.763. Oppenheimer and Carlson termed the latter neutrino, Pauli's neutron, the 'magnetic neutron' so as to distinguish it from the neutron which was at that time associated with the penetrating beryllium radiation. This particle, our neutron, had been announced by Chadwick about six months earlier to be a constituent of the nucleus, with a mass very close to the proton's.

In the summer of 1932, there was still confusion about the two 'neutrons'. Oppenheimer and Carlson thus thought that Pauli's magnetic neutron might well have a mass like the one indicated in Chadwick's experiment. As to Pauli's wave equation for the magnetic neutron -the neutrino- it was first given at a seminar in Ann Arbor, USA, in the summer of 1931, but was not published. According to Oppenheimer and Carlson, it was identical to the Dirac equation as long as no field was present; in the presence of an electromagnetic field it was obtained by adding a term including the magnetic moment of the neutrino. Pauli's unpublished equation was a four-component equation, preserving left-right symmetry.

19. The content of Engels' reflections on the dialectics of nature, and the position of monopoles herein, dates from about 1878 but was only published in 1925. Engels claimed that the dialectical laws preclude the existence of magnetic monopoles. See [56], pp.61-62.

20. Only five years earlier, Oskar Klein had suggested an explanation of the quantization of electricity in terms of five-dimensional relativity quantum theory. See [111]. And before that time, there were several attempts to deduce the electrical charge from Einstein's and Weyl's gravitation theories.

21. When Pauli and Weisskopf in 1934 quantized the Klein-Gordon equation, and showed that it describes hypothetical particles obeying Bose-Einstein statistics, they joculary paraphrased Dirac in asking why "Die Natur ... keinen Gebrauch gemacht hat" of negatively charged bosons of spin zero ([139],p.713). Only much later it became known that nature does, in fact, 'make use' of these particles, as the Pauli-Weisskopf theory applies to pions.

22. This number is based on an examination of the abstracts appearing in the Physikalische Berichte. The papers are from 1931(2), 1935(2) and 1938(1).

23. See [52]. The Ehrenhaft-Dirac correspondence is not included in the AHQP material. The only full description of Ehrenhaft's work on magnetic charges appeared in a non-physics journal, the Philosophy of Science([68']). In this paper one may find further references to his work on the subject. It is remarkable that Ehrenhaft did not mention Dirac's theory but presented his claim for magnetic poles in a purely empirical way. For a sympathetic portrait of Ehrenhaft in his later days, see Feyerabend ([76],p.109) who attended Ehrenhaft's lectures in 1947. We learn from Feyerabend that Ehrenhaft rejected not only the elementary electron and the Maxwell equation $\nabla \cdot \vec{B}=0$, but also relativity and quantum mechanics. No wonder that Dirac did not like to get associated with the unorthodox Viennese physicist! Ehrenhaft died in 1952, and thus did not live to experience the renewed interest in monopoles and fractionally charged particles (quarks).

24. See [145]. Price interpreted the event, an ionizing track in a plastic emulsion, as being due to a monopole with a strength corresponding to 137e, velocity at about c/2, and mass greater than 200 proton masses. A summary of earlier experimental attempts to detect the Dirac pole may be found in [2].

25. According to [56],p.49 and to New Scientist, 21 August 1975, p.412.

26. Transcript of interview deposited at the Niels Bohr Institute, Copenhagen. Cf. [115].

27. This interpretation rested on the basic assumption that

the electromagnetic four-potential A_μ satisfies the condition $A_\mu A^\mu = k^2$, k being a universal constant which by hypothesis was taken to be m/e. Observing that $k^{-1}A_\mu = (e/m)A_\mu$ has the dimension of a four-velocity, Dirac argued that it would be the 'aether velocity'.

28. See [151]. Dirac, in his 1948 paper, formulated monopole electrodynamics in an action principle but had to impose the constraint that nodal lines can never pass through charged particles. Dirac's theory was not, therefore, given by a pure action principle.

29. This conclusion was not accepted by all physicists. Mendel Sachs, an authority on relativity theory, thus disagreed with Bilanuik and Sudarshan as to whether tachyons were allowed by relativity: "... if such particles should be found, I should have to conclude ... that the theory of relativity would have been refuted" ([152],p.48).

30. For a clear survey of the main points in Milne's theory, see his "Gravitation Without General Relativity", pp. 409-436 in [153].

31. In his work "The Sand Reckoner", written about 220 BC, Archimedes computed the maximum number of grains of sand the universe could contain, to circa 10^{63}. Each of Archimedes' grains of sand may be estimated to contain 10^{17} nucleons, hence the magic number 10^{80}! See [89].

32. The unsatisfactory numerical relationship between the age of the universe and the Hubble constant, as in Dirac's theory, caused discredit to Big-Bang theories at the end of the thirties. The inconsistencies are not real, however, and need not to be rescued by Dirac's artificial theory of time-dependent radioactive decay laws. It disappears with the later accepted values of Hubble's con-

stant, which are smaller by a factor ten to what was thought in 1937.

33. Jordan's occupation with cosmological theory à la Dirac started before the war [107]. Under the impact of Eddington's and Dirac's theories, he engaged in the very same type of numerological reasoning. In accordance with his positivistic conception of science, Jordan, however, emphasized that there was nothing speculative about such attempts. The various relationships between the constants of nature were, Jordan stressed, "hypothesenfreie blosse Umrechnungen der Erfahrungstatsachen" (ibid.,p.515). Far from being speculative cosmology, Jordan termed his development of Dirac's and Eddington's systems for empirical cosmology [108].

34. Haldane praised the works of Milne and Dirac for having "introduced the historical process into exact physics". See [87],p.76.

35. To be sure, Einstein's field equations may be written in neat forms, such as $G_{\nu\mu} = 0$ or $G_{\nu\mu} = \lambda g_{\nu\mu}$, appearing to be even simpler formulae than Newton's law. The simplicity is, however, a trick played by compact notation. Behind the innocent-looking tensor quantities, conceptual and technical complications are disguised. Cf. also that all the laws of physics can be contained in one grand 'simple' law, U=0, such as shown by Feynman, not without irony. See [77],p. 25:10. In these cases, as in axiomatization attempts in general, the obtained simplicity is illusory.

36. In a talk on "Relativity and Quantum Mechanics", given in Austin, Texas, in 1970. Quoted from [122],p.59.

37. See [8]. Born argued forcefully against Eddingtonian ideas of pure theoretical reasoning and for the value of experiments and the inductive method even in the most abstract theories. Einstein's general theory of relativity was by

Born considered to be "a gigantic synthesis of a long chain of empirical results, not a spontaneous wave brain" (p.14). And for quantum mechanics, "it was ... an essentially inductive line of reasoning which led to the most abstract theory known in physics" (p.20). But neither Born was always immune to the intellectual magic of Einstein's theory. In 1920, he saluted it for its "grandeur, the boldness, and the directness of the thought", which made the world-picture "more beautiful and grander" [6].

38. That is, in his work with general relativity and later on. For the much discussed question of whether the young Einstein followed a positivistic method or not, see [133] and [84]. See also the contributions in [153].

39. When Einstein received the news of the measurements of the eclipse expedition he said, "But I knew that the theory is correct." And on the question, what if the measurements had disagreed with the theory, Einstein countered, "Then I would have been sorry for the dear Lord - the theory is correct." See [98],p.236. Einstein's reaction to an earlier eclipse measurement in 1914, which indicated a discrepancy between theory and observation, was similar. He refused to accept the experimental results as a disproof of the theory. It is ironical, then, that Einstein's theory and the 1919 experiment served as a decisive inspiration for the young Karl Popper in the creation of his methodology of science. Certainly, Einstein does not fit into falsificationism. On the other hand, Einstein's reactions to the experimental tests of relativity are somewhat perplexing to the historians of science. For in reply to Eddington's letter, announcing the successful results of the eclipse observations, Einstein stressed the importance of another test, viz. the gravitational displacement of solar absorption lines towards the red. Concerning this test, Einstein told Eddington that, "If it were proved that this effect does not exist in nature,

then the whole theory would have to be abandoned" (letter from Einstein to Eddington, 15 December 1919. Quoted in translation from [60],p.41). This attitude of course fits perfectly well to Popper's views.

40. In 1906 Kaufmann performed careful experiments on the mass of moving electrons and announced his experimental results to be a categorical disproof of Einstein's theory. See [98],p.235.

41. In a conversation with Freeman Dyson, see [16],p.27.

42. In a letter to V.Weisskopf, 17 January 1957. Quoted from [80],p.215.

43. For a useful review of the dispute, see [65].

44. This is not to say that socialists, or authors from the communist bloc, have been immune to the quest for beauty in theoretical science (and why should they?). For a Russian appraisal of the beauty associated with Einstein's theory, see [155]. Cf. also that Landau and Lifschitz, in their very technical nine-volume course in theoretical physics, only expressed any display of emotion at one place. That was in connection with "The theory of gravitational fields ... established by Einstein (and finally formulated by him in 1916) and represents probably the most beautiful of all existing theories. It is remarkable that it was developed by Einstein in a purely deductive manner and only later was substantiated by astronomical observations." ([116],p.227)

45. For Heisenberg's methodology and philosophy of science, see further [90] and [100].

46. "I admired Bohr very much. He seemed to be the deepest thinker I ever met" ([53],p.134). "While I was very much

impressed by what Bohr said, his arguments were mainly of a qualitative nature ... What I wanted was statements which could be expressed in terms of equations, and Bohr's work very seldom provided such statements." (ibid.,p.116)

47. Keats in a letter to B.Bailey, 22 November 1817. Quoted from the <u>Norton Anthology of English Literature</u>, vol.2, New York 1968, p.569.

48. This example is due to Léon Rosenfeld in [69],p.38.

-oOo-

LITERATURE CITED.

[1] Agassi, J. *Science in Flux* (Boston Studies in the Philosophy of Science vol.28), Dordrecht 1975.

[2] Amaldi, E. and Cabibbo, N. "On the Dirac Magnetic Poles" pp.183-212 in Salam and Wigner [122].

[3] Bachelard, G. *Nej'ets Filosofi*, Copenhagen 1977 (La Philosophie du non, Paris 1940).

[4] Bernal, J.D. *Science in History*, 4 vols., London 1969.

[5] Bilanuik, O.-M. and Sudarshan, E.C. "Particles Beyond the Light Barrier," *Physics Today*, May 1969, 43-51.

[6] Born, M. *Einstein's Theory of Relativity*, London 1924 (German original 1920).

[7] Born, M. "Some Philosophical Aspects of Modern Physics," *Proc.Roy.Soc.(Edinburgh)*, *57* (1936/37), 1-18.

[8] Born, M. *Experiment and Theory in Physics*, Cambridge 1943.

[9] Born, M. *Physik im Wandel meiner Zeit*, Braunschweig 1959.

[10] Broad, C.D. "The Philosophical Aspects of the Theory of Relativity," *Mind*, *29* (1920), 430-437.

[11] Bromberg, J. "The Impact of the Neutron: Bohr and Heisenberg," *Hist.Stud.Phys.Sci.*, *3*(1971), 307-343.

[12] Bromberg, J. "Remarks on Hole Theory and Dirac's Methodology," *Proceed. 14th Int. Congress Hist.Sci.*, 1975, vol.2, 233-236.

[13] Bromberg, J. "The Concept of Particle Creation Before and After Quantum Physics," *Hist.Stud. Phys.Sci.*, *7*(1976), 161-183.

[14] Bronowski, J. *Science and Human Values*, London 1964.

[15] Bunge, M. "The Weight of Simplicity in the Construction and Assaying of Scientific Theories," *Phil.Sci.*, *28*(1961), 120-149.

[16] Chandrasekhar,S. "Beauty and the Quest for Beauty in Science," Physics Today, July 1979, 25-30.

[17] Chedd,G. "Monopole Marvels," New Scientist,11 September 1975, p.607.

[18] Clay,R.W. and Crouch,P.C. "Possible Observation of Tachyons Associated with Extensive Air Showers," Nature, 248(1974), 28-30.

[19] Darwin,C.G. "The Wave Equation of the Electron," Proc.Roy.Soc.(London), A118(1928), 654-680.

[20] Dicke,R.H. "Dirac's Cosmology and Mach's Principle," (with Dirac's reply) Nature, 192(1961), 440-441.

[21] Dingle,H. "Modern Aristotelianism," Nature, 139 (1937), 784-786.

[22] Dingle,H. " Deductive and Inductive Methods in Science: a Reply," Nature, 139(1937), 1011-1012.

[23] Dirac,P.A.M. "Note on the Relativity Dynamics of a Particle," Phil.Mag., 47(1924), 1158-1159.

[24] Dirac,P.A.M. "The Physical Interpretation of the Quantum Dynamics," Proc.Roy.Soc.(London), A113(1927), 621-641.

[25] Dirac,P.A.M. "The Quantum Theory of the Electron, I," Proc.Roy.Soc.(London), A117(1928), 610-624.

[26] Dirac,P.A.M. "A Theory of Electrons and Protons," Proc. Roy.Soc.(London), A126(1930), 360-365.

[27] Dirac,P.A.M. "On the Annihilation of Electrons and Protons," Proc.Cambridge Phil.Soc., 26 (1930), 361-375.

[28] Dirac,P.A.M. "The Proton," Nature, 126(1930), 605.

[29] Dirac,P.A.M. The Principles of Quantum Mechanics, Oxford 1930.

[30] Dirac,P.A.M. "Quantized Singularities in the Electromagnetic Field," Proc.Roy.Soc.(London), A 133(1931), 60-72.

[31] Dirac,P.A.M. "Theory of Electrons and Positrons," pp. 320-325 in Nobel Lectures, Physics 1922-41, Amsterdam 1965.

[31'] Dirac,P.A.M. "Discussion of the Infinite Distribution of Electrons in the Theory of the Positron," Proc.Cambridge Phil.Soc., 30 (1934), 150-163.

[32] Dirac,P.A.M. "Relativistic Wave Equations," Proc.Roy. Soc.(London), A 155(1936), 447-459.

[33] Dirac,P.A.M. "Does Conservation of Energy Hold in Atomic Processes?" Nature, 137(1936), 298-299.

[34] Dirac,P.A.M. "The Cosmological Constants," Nature, 139 (1937), 323.

[35] Dirac,P.A.M. "Complex Variables in Quantum Mechanics," Proc.Roy.Soc.(London), A 160(1937), 48-59.

[36] Dirac,P.A.M. "A New Basis for Cosmology," Proc.Roy. Soc.(London), A 165(1938), 199-208.

[37] Dirac,P.A.M. "Classical Theory of Radiating Electrons," Proc.Roy.Soc.(London), A 167(1938), 148-169.

[38] Dirac,P.A.M. "The Relation Between Mathematics and Physics," Proc.Roy.Soc.(Edinburgh), 59 (1939), 122-129.

[39] Dirac,P.A.M., Peierls,R. and Pryce,M. "On Lorentz Invariance in the Quantum Theory," Proc. Cambridge Phil.Soc., 38(1942), 193-200.

[40] Dirac,P.A.M. "The Physical Interpretation of Quantum Mechanics," Proc.Roy.Soc.(London), A 180 (1942), 1-39.

[41] Dirac,P.A.M. "The Theory of Magnetic Poles," Phys.Rev., 74(1948), 817-830.

[42] Dirac,P.A.M. "Is there an Aether?" Nature, 168(1951), 906-907.

[43] Dirac,P.A.M. "A New Classical Theory of Electrons," Proc.Roy.Soc.(London), A 209(1951), 291-296.

[44] Dirac,P.A.M. "Quantum Mechanics and the Aether," Sci. Monthly, 78(1954), 142-146.

[45] Dirac,P.A.M. "The Evolution of the Physicist's Picture of Nature," Sci.Amer., 208 no.5(1963), 45-53.

[46] Dirac,P.A.M. "Quantum Electrodynamics without Dead Wood," Phys.Rev., 139(1965), 684-690.

[47] Dirac,P.A.M. Lectures on Quantum Field Theory, New York 1966.

[48] Dirac,P.A.M. "Methods in Theoretical Physics," pp.21-28 in From a Life of Physics, IAEA Bulletin, Vienna 1969.

[49] Dirac,P.A.M. "Can Equations of Motion be Used in High Energy Physics?" Physics Today, 23 no.4, (1970), 29-31.

[50] Dirac,P.A.M. "A Positive Energy Relativistic Wave Equation," Proc.Roy.Soc.(London), A 322 (1971), 435-445.

[51] Dirac,P.A.M. The Development of Quantum Theory, New York 1971.

[52] Dirac,P.A.M. "Ehrenhaft, the Subelectron and the Quark," pp.290-293 in [163].

[53] Dirac,P.A.M. "Recollections of an Exciting Era," pp. 109-146 in [163].

[54] Dirac,P.A.M. "Long Range Forces and Broken Symmetries," Proc.Roy.Soc.(London), A 333(1973), 403-418.

[55] Dirac,P.A.M. "Cosmological Models and the Large Number Hypothesis," Proc.Roy.Soc.(London), A 338(1974), 439-446.

[56] Dirac,P.A.M. Directions in Physics, New York 1975.

[57] Dirac,P.A.M. "The Mathematical Foundations of Quantum Theory," pp.1-18 in A.R.Marlow(ed.), Mathematical Foundations of Quantum Theory, New York 1978.

[58] Dirac,P.A.M. "The Test of Time," The Unesco Courier, May 1979, 17-23.

[59] Dirac,P.A.M. "The Excellence of Einstein's Theory of Gravitation," Impact of Science on Society, 29(1979), 11-14.

[60] Douglas,A.V. The Life of Arthur Stanley Eddington, London 1956.

[61] Drell,S.D. "When is a Particle?" Physics Today, June 1978, 23-32.

[62] Dyson,F.J. "The Fundamental Constants and their Time Variation," pp.213-236 in [122].

[63] Eddington,S.A. Space, Time, and Gravitation, Cambridge 1920.

[64] Eddington,S.A. The Nature of the Physical World, Cambridge 1928.

[65] Eddington,S.A. Relativity Theory of Protons and Electrons, Cambridge 1936.

[66] Eddington,S.A. New Pathways in Science, Cambridge 1935.

[67] Einstein,A. "On the Method of Theoretical Physics," in The World as I see it, London 1935.

[68] Ehrenhaft,F. "New Evidence for the Magnetic Current," Phys.Rev., 67(1945), 63.

[68'] Ehrenhaft,F. and Banet,L. "Is there 'True Magnetism' or Not?" Phil.Sci., 8(1941), 458-462.

[69] Elkana,Y.(ed.) The Interaction Between Science and Philosophy, Atlantic Highlands (N.J.), 1974.

[70] Elsasser,W.M. Memoirs of a Physicist in the Atomic Age, New York 1978.

[71] Engels,F. Dialektik der Natur, Berlin(DDR) 1973.

[72] Feather,N. "A History of Neutrons and Nuclei," Contemporary Physics, 1(1960),191-203 and 257-266.

[73] Feinberg,G. "Possibility of Faster-Than-Light Particles," Phys.Rev., 159(1967), 1089-1105.

[74] Feuer,L.S. Einstein and the Generations of Science, New York 1974.

[75] Feuer,L.S. "Teleological Principles in Science," Inquiery, 21(1977), 377-407.

[76] Feyerabend,P. Science in a Free Society, London 1978.

[77] Feynman,R.P. The Feynman Lectures on Physics, vol.2, Reading(Mass.), 1964.

[78] Ford,K.W. "Magnetic Monopoles," Sci.Amer., 209 (December 1963), 122-131.

[79] Forman, P. "Weimar Culture, Causality, and the Quantum Theory, 1918-1927: Adaption by German Physicists and Mathematicians to a Hostile Intellectual Environment," Hist. Stud. Phys. Sci., 3 (1971), 1-115.

[80] Franklin, A. "The Discovery and Nondiscovery of Parity Nonconservation," Stud. Hist. Phil. Sci., 10 (1979), 201-257.

[81] Galilei, G. Dialogues Concerning the Two Great Systems of the World, Los Angeles 1953.

[82] Gamow, G. Tredive År der Rystede Fysikken, Copenhagen 1968 (Thirty Years that Shook Physics, 1966).

[83] Goldberg, S. "Max Planck's Philosophy of Nature and his Elaboration of the Special Theory of Relativity," Hist. Stud. Phys. Sci., 7 (1976), 125-161.

[84] Gutting, G. "Einstein's Discovery of Special Relativity," Phil. Sci., 39 (1972), 51-68.

[85] Haas, A. "An Attempt to a Purely Theoretical Derivation of the Mass of the Universe," Phys. Rev., 49 (1936), 411-412.

[86] Hadamard, J. The Psychology of Invention in the Mathematical Field, New York 1954 (first edition 1945).

[87] Haldane, J.B.S. The Marxist Philosophy and the Sciences, New York 1939.

[88] Hanson, N.R. The Concept of the Positron, Cambridge 1963.

[89] Harrison, E.R. "The Cosmic Numbers," Physics Today, December 1972, 30-34.

[90] Heelan, P.A. "Heisenberg and Radical Theoretic Change," Zs. Allgemeine Wissenschaftstheorie, 6 (1975), 113-136.

[91] Heisenberg, W. "Über den anschaulichen Inhalt der Quantentheoretischen Kinematik und Mekanik," Zs. f. Phys., 43 (1927), 172-198.

[92] Heisenberg, W. "Quantum Theory and its Interpretation," in S. Rozental (ed.), Niels Bohr, Amsterdam 1967.

[93] Heisenberg, W. Del og Helhed, Copenhagen 1971 (Der Teil und das Ganze, Munich 1969).

[94] Heisenberg,W. "Development of Concepts in the History of Quantum Theory," pp.264-275 in [123].

[95] Holton,G. "Themata in Scientific Thought," pp.3-24 in Holton, The Scientific Imagination: Case Studies, Cambridge(Mass.) 1978.

[96] Holton,G. "Subelectrons, Presuppositions, and the Millikan-Ehrenhaft Dispute," pp.25-83 in [72].

[97] Holton,G. "Einstein and the 'Crucial' Experiment," Am.J.Phys., 37(1969), 968-982.

[98] Holton,G. Thematic Origins of Scientific Thought, Cambridge(Mass.), 1973.

[99] Huntley,H.E. The Divine Proportion. A Study in Mathematical Beauty, New York 1970.

[100] Hörz,H. Werner Heisenberg und die Philosophie, Berlin(DDR) 1966.

[101] Jammer,M. The Conceptual Development of Quantum Mechanics, New York 1966.

[102] Jeans,J. The Mysterious Universe, London 1930.

[103] Jeans,J. The New Background of Science, Cambridge 1933.

[104] Jones,L.W. "A Review of Quark Search Experiments," Rev.Mod.Phys., 49(1977), 717-752.

[105] Jordan,P. "Zur Quantenelektrodynamik,III. Eichinvariante Quantelung und Diracsche Magnetpole," Zs.f.Phys., 97(1935),535-537.

[106] Jordan,P. Anschauliche Quantentheorie, Berlin 1936.

[107] Jordan,P. "Die physikalischen Weltkonstante," Die Naturwissenschaften, 25(1937),513-517.

[108] Jordan,P. "Zur empirischen Kosmologie," Die Naturwissenschaften, 26(1938),417-421.

[109] Jordan,P. Schwerkraft und Weltall, Braunschweig 1952.

[110] Kepler,R. Astronomi Opera Omnia, vol.6 (ed.C.Frisch 1858)

[111] Klein,O. "The Atomicity of Electricity as a Quantum Law," Nature, 118(1926), 516.

[112] Kline,M. Mathematics in Western Culture, London 1972.

[113] Kragh,H. "The Genesis of Dirac's Relativity Theory of Electrons," Arch.Hist.Ex.Sci., 1980 (fortcoming).

[114] Kuhn,T.S. The Structure of Scientific Revolutions, Chicago 1962.

[115] Kuhn,T.S. et al. Sources for History of Quantum Physics, Philadelphia 1967.

[116] Landau,L.D. and Lifschitz,E.M. The Classical Theory of Fields, Oxford 1971.

[117] Lovejoy,A.O. The Great Chain of Being, Cambridge(Mass.) 1976 (first edition 1936).

[118] Lyon,H.T.(ed.) Keats' Well-Read Urn, London 1958.

[119] Mandelstam,S. and Yourgraw,W. Variational Principles in Dynamics and Quantum Theory, London 1958.

[120] Maxwell,J.C. Treatise on Electricity and Magnetism, third edition, Oxford 1891.

[121] McCusker,C.B. et al. "Nuclear Interactions and Cosmic Radiation Around 10^6 GeV," Phys.Rev., 177 (1969), 1902-1920.

[122] Mehra,J. "'The Golden Age of Theoretical Physics': P.A.M.Dirac's Scientific Work from 1924 to 1933," pp.17-60 in A.Salam and E.P.Wigner (eds.) Aspects of Quantum Theory, Cambridge 1972.

[123] Mehra,J.(ed.) The Physicist's Conception of Nature, Dordrecht 1973.

[124] Merleau-Ponty,J. Philosophie et théorie physique chez Eddington, Paris 1965.

[125] Mignani,R. "Strong Black Holes and Quark-Monopole Identification," Lett. Nuovo Cimento, 16 (1976),6-8.

[126] Miller,A. "Visualization Lost and Regained: The Genesis of the Quantum Theory in the Period 1913-27," pp.74-102 in J.Wechsler(ed.), On Aesthetics in Science, Cambridge(Mass.) 1978.

[127] Miller,R.W. "Truth in Beauty," Amer.Phil.Quarterly, 16(1979), 317-325.

[128] Mills,J.S. A System of Logic, 2 vols., London 1843.

[129] Milne,E.A. "On the Origin of Laws of Nature," Nature, 139(1937),997-999.

[130] Milne,E.A. Relativity, Gravitation and World-Structure, Oxford 1935.

[131] Mituo,T. "Methodological Approaches in the Development of the Meson Theory of Yukawa in Japan," pp.24-38 in S.Nakayama (ed.), Science and Society in Modern Japan, Tokyo 1974.

[132] (Nature) "Physical Science and Philosophy," Nature, 139(1937), 1000-1012.

[133] Neidorf,R. "Is Einstein a Positivist?" Phil.Sci., 30 (1963),173-188.

[134] Newton,I. Principia Mathematica, London 1687 (The Mathematical Principles of Nature / The System of the World, ed. F.Cajori, Berkeley 1946).

[135] Oppenheimer,J.R. "Two Notes on the Probability of Radiative Transitions," Phys.Rev., 35(1930), 939-947.

[136] Oppenheimer,J.R. "On the Theory of Electrons and Protons," Phys.Rev., 35(1930), 562.

[137] Oppenheimer,J.R. and Carlson,J.F. "The Impacts of Fast Electrons and Magnetic Neutrons," Phys.Rev., 41(1932), 763-792.

[138] Pauli,W. "Die allgemeinen Prinzipien der Wellenmechanik," in Handbuch der Physik, band 24/I, Berlin 1933.

[139] Pauli,W. and Weisskopf,V. "Über die Quantisierung der skalaren relativistischen Wellengleichung," Helv.Phys.Acta, 7(1934), 709-731.

[140] Poincaré,H. "Remarques sur une expérience de M.Birkeland," Comptes Rendus, 123(1896), 530-533.

[141] Poincaré,H. Science and Method,New York 1960 (first edition Paris 1908).

[142] Polanyi,M. Personal Knowledge, Chicago 1958.

[143] Polanyi,M. "Beauty, Elegance, and Reality in Science" pp.102-106 in S.Körner(ed.), Observation and Interpretation in the Philosophy of Physics, New York 1962.

[144] Popper,K.R. The Logic of Scientific Discovery, London 1958.

[145] Price,P.B. et al. "Evidence for Detection of a Moving Magnetic Monopole," Phys.Rev.Letters, 35 (1975), 487-490.

[146] Pyenson,L. "Hermann Minkowksi and Einstein's Special Theory of Relativity," Arch.Hist.Ex.Sci., 17(1977), 71-95.

[147] Pyenson,L. "Physics in the Shadow of Mathematics: The Göttingen Electron-Theory Seminar of 1905," Arch.Hist.Ex.Sci., 21(1979), 55-89.

[148] Recami,E. and Mignani,R. "A New Experimental and Theoretical Outlook on Magnetic Monopoles," pp.321-324 in W.C.Price and S.S.Chiddick (eds.), The Uncertainty Principle and Foundations of Quantum Mechanics, London 1977.

[149] Richter,S. Wolfgang Pauli, Aurau 1979.

[150] Rosenbaum,D. Physics Today, December 1976, 47.

[151] Rosenbaum,D. "Proof of the Impossibility of a Classical Action Principle for Magnetic Monopoles and Charges without Subsidiary Conditions," Phys.Rev., 147(1966), 891-895.

[152] Sachs,M. "More about Tachyons," Physics Today, December 1969, 48.

[153] Schilpp,P.A.(ed.) Albert Einstein: Philosopher-Scientist, Evanston(Ill.), 1949.

[154] Schrödinger,E. Science, Theory and Man, New York 1957.

[155] Smilga,V. In the Search for Beauty, Moscow 1970.

[156] Sommerfeld,A. "Vereenvoudigte afleidung van het veld van en de krachten werkende op een electron bij willekeurige bewegung," Amst.Akad., 13 (1904),431-452 (=Gesammelte Schriften, vol. 2, pp.17-38).

[157] Sommerfeld, A. "Zur Elektronentheorie III. Ueber Lichtgeschwindigkeits- und Ueberlichtgeschwindigkeits-Elektronen," Nachrichten Kgl. Gesellschaft Wiss. Göttingen Math.-Phys. Klasse, 1905, 201-235 (=Gesammelte Schriften, vol.2, pp.148-182).

[158] Tamm, I. "Die verallgemeinerten Kugelfunktionen und die Wellenfunktionen eines Elektrons im Feld eines Magnetpoles," Zs.f.Phys., 71(1931), 141-150.

[159] Thomson, J.J. Elements of the Mathematical Theory of Electricity and Magnetism, Cambridge (Mass.). 1900.

[160] Tipler, F.J. "Direct-Action Electrodynamics and Magnetic Monopoles," Nuovo Cimento, 28(1975) 446-452.

[161] Weinberg, S. "Reflections of a Working Scientist," Dædalus, 103(Summer 1972).

[162] Weinberg, S. "The Search for Unity: Notes for a History of Quantum Field Theory," Dædalus, 108 (Fall 1977), 17-34.

[163] Weiner, C. (ed.) History of Twentieth Century Physics, New York 1977.

[164] Werskey, P.G. "British Scientists and 'Outsider' Politics 1931-1945," pp.231-250 in B.Barnes (ed.), Sociology of Science, London 1972.

[165] Werskey, P.G. The Visible College, New York 1978.

[166] Weyl, H. "Elektron und Gravitation," Zs.f.Phys., 56(1929), 330-352.

[167] Weyl, H. Gruppentheorie und Quantenmechanik, Berlin 1931.

[168] Weyl, H. The Open World. Three Lectures on the Metaphysical Implications of Science, Yale 1932.

[169] Whittaker, E. From Euclid to Eddington, Cambridge 1949.

[170] Wigner, E.P. Symmetries and Reflections, Bloomington (Ind.) 1967.

[171] Winterberg,F. "Quarks, Magnetic Monopoles and Negative Mass," Lett.Nuovo Cimento, 13(1975), 697-703.

[172] Witt-Hansen,J. Exposition and Critique of the Conceptions of Eddington Concerning the Philosophy of Physical Science, Copenhagen 1958.

[173] Yukawa,H. "On the Interaction of Elementary Particles, I," Proc.Phys.-Math.Soc.Japan, 17 (1935), 48-57.

[174] Margenau,H. "Methodology of Modern Physics," pp.52-89 in Physics and Philosophy: Selected Essays, Dordrecht 1978 (first published in Phil.Sci., 2(1935), 48-72 and 164-187).

-o8o-

16/79 "STATISTIK I KRÆFTFORSKNINGEN". Projektrapport af Michael Olsen og Jørn Jensen. Vejleder: Jørgen Larsen.

17/79 "AT SPØRGE OG AT SVARE i fysikundervisningen" Albert Christian Paulsen.

18/79 "MATHEMATICS AND THE REAL WORLD", Proceedings of an International Workshop, Roskilde university centre (Denmark), 1978. Preprint. Bernhelm Booss & Mogens Niss (eds.).

19/79 "GEOMETRI, SKOLE OG VIRKELIGHED". Projektrapport af Tom J. Andersen, Tommy R. Andersen og Per H.H. Larsen. Vejleder: Mogens Niss.

20/79 "STATISTISKE MODELLER TIL BESTEMMELSE AF SIKRE DOSER FOR CARSINOGENE STOFFER". Projektrapport af Michael Olsen og Jørn Jensen. Vejleder: Jørgen Larsen.

21/79 "KONTROL I GYMNASIET - FORMÅL OG KONSEKVENSER". Projektrapport af Crilles Bacher, Per S. Jensen, Preben Jensen og Torben Nysteen.

22/79 "SEMIOTIK OG SYSTEMEGENSKABER (1)". 1-port lineært response og støj i fysikken. Peder Voetmann Christiansen.

23/79 "ON THE HISTORY OF EARLY WAVE MECHANICS - with special emphasis on the role of relativity". Helge Kragh.

24a/80 "MATEMATIKOPFATTELSER HOS 2.G'ERE" 1. En analyse.
24b/80 "MATEMATIKOPFFATTELSE HOS 2.G'ERE" 2. Interviewmateriale. Projektrapport af Jan Christensen og Knud Lindhardt Rasmussen. Vejleder: Mogens Niss.

25/80 "EKSAMENSOPGAVER" Dybdemodulet/fysik 1974-79.

26/80 "OM MATEMATISKE MODELLER". En projektrapport og to artikler. Jens Højgaard Jensen m.fj.

27/80 "METHODOLOGY AND PHILOSOPHY OF SCIENCE IN PAUL DIRAC's PHYSICS" Helge Kragh.

ISSN 0106 - 6242