Quantum Gravity in the First Half of the Twentieth Century

A Sourcebook

Edition Open Access

The Edition Open Access (EOA) platform was founded to bring together publication initiatives seeking to disseminate the results of scholarly work in a format that combines traditional publications with the digital medium. It currently hosts the open-access publications of the "Max Planck Research Library for the History and Development of Knowledge" (MPRL) and "Edition Open Sources" (EOS). EOA is open to host other open access initiatives similar in conception and spirit, in accordance with the *Berlin Declaration on Open Access to Knowledge* in the sciences and humanities, which was launched by the Max Planck Society in 2003.

By combining the advantages of traditional publications and the digital medium, the platform offers a new way of publishing research and of studying historical topics or current issues in relation to primary materials that are otherwise not easily available. The volumes are available both as printed books and as online open access publications. They are directed at scholars and students of various disciplines, as well as at a broader public interested in how science shapes our world.

Edition Open Access
2018

Quantum Gravity in the First Half of the Twentieth Century
A Sourcebook

Alexander S. Blum and Dean Rickles (eds.)

Sources 10

Edition Open Sources

Edition Open Sources (EOS) pioneers a new paradigm in the publishing of historical sources. Academic editions of primary sources in the history of science are published in online, digital, and print formats that present facsimiles, transcriptions, and often translations of original works with an introduction to the author, the text, and the context in which it was written. The sources are historical books, manuscripts, documents, or other archival materials that are otherwise difficult to access. EOS is a cooperation between the University of Oklahoma Libraries, the Department for the History of Science der University of Oklahoma, and the Max Planck Institute for the History of Science.

Sources 10

Submitters: Carlo Rovelli and Jürgen Renn

Cover: A selection of title pages from the source texts reproduced in this volume.

Images: The images in this volume have been prepared for publication by the Digitization Group of the Library of the Max Planck Institute for the History of Science.

ISBN 978-3-945561-31-7
First published 2018 by Edition Open Access
Max Planck Institute for the History of Science
http://www.edition-open-access.de
Printed and distributed by
PRO BUSINESS digital printing Deutschland GmbH, Berlin

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available at http://dnb.d-nb.de.

Contents

Preface

Bryce DeWitt recalls having a conversation with Pauli in 1949, at the Institute for Advanced Study, in which Pauli asked him what he was working on,[1] to which DeWitt responded: "trying to quantize the gravitational field." Pauli was something of a veteran of quantum gravity by this time. His response, after shaking and nodding his head a few times ("die Paulibewegung"), was: "That is a very important problem—but it will take someone really smart!".[2] More than half a century since this conversation, and with the work of many genuinely "really smart" people (including several geniuses), the problem has, of course, not yet been resolved. Yet still the field has accumulated a rich and interesting past that has yet to be properly studied. It has become customary to mark a certain stage of maturity of a theory by producing a volume of sources of early papers from which that maturity emerged.

Though quantum gravity has not yet achieved full *scientific* maturity, it has at least achieved chronological maturity, with almost a century of struggle behind it. Therefore, we feel it is entirely appropriate to treat this old timer with some respect, of which it has not received all that much from the history of physics. A volume providing a historical overview, after so long without one, can be beneficial to the current and future generation of physicists working on the problem, in order to see how far research on the problem (and the way the very problem itself is conceptualized) has come—this further provides a fresh perspective on what still remains to be done. It might point to further refinements of how we understand the problem so that it can finally be resolved.

As Julian Schwinger pointed out in the preface to his own collection of papers from the history of quantum electrodynamics,[3] any such selection of sources is bound to reflect the particular viewpoint of the editor(s). Following Schwinger, we briefly describe our "selection process." Despite the fact that the period we cover spans only 35 years, it was necessary to be fairly brutal in rejecting papers for which a case for inclusion could easily be made. Likewise, it is likely that cases could be made for *excluding* many of the papers we decided to include. The point is, the sources chosen are an imperfect reflection of the development of a field, and one important reason for this is that "the field" in question has always been somewhat slippery and hard to define, but especially so in its earliest phases of development. One can't, for example, point to particular phenomena that the theory will describe since any such phenomena would be experimentally and observationally very remote. Moreover, in the earliest phases of research, the ingredient theories (general relativity and quantum mechanics) were themselves still being worked through and, in the case of the latter, were not properly formulated for some time (as Part I indicates). Thus one finds the definition of the problem of quantum gravity is non-stationary on account of being largely at the mercy of wider developments in quantum mechanics and general relativity in our chosen time period—one sees this especially clearly in Part II, but it is really a general feature.

---

[1] DeWitt (then still using the name Carl Bryce Seligman) had only just finished his doctoral thesis on quantum gravity (under Julian Schwinger at Harvard: submitted in December, 1949), and was interested in the possibility of a postdoc at ETH.

[2] "Quantum Gravity: Yesterday and Today." *General Relativity and Gravitation* **41**, 2009: p. 414.

[3] *Selected Papers on Quantum Electrodynamics* (Dover, 1958).

Firstly, let us explain the period we have restricted our sources to, namely <1950. This is a relatively short snapshot of history, but it has the advantage of revealing the steps taken before the rather dramatic explosion of work in the 1950s—this explosion was due to a variety of factors beyond internal advances in physics, including the emergence of new schools of research in general relativity (especially Peter Bergmann's, Hermann Bondi's, Alfred Schild's, Leopold Infeld's, and John Wheeler's) and new sources of funding (especially the NSF, the ONR, and other military and philanthropic programmes).[4] We also find that the pre-1950 research already includes many of the main lines of attack and the main general arguments for (and against) quantization. Focusing in on the very earliest period sharpens the physical intuitions behind the various choices (of formalism, terminology, and more) that have since been assimilated or forgotten.

We will now go on to explain why there are papers in here (many, in fact) that are not strictly "quantum gravitational". One can quite usefully think of the development of research programmes (and questions) in a field in terms of "evolutionary trees." Pursuing this in the case of organisms eventually leads one outside of the species of interest. Or one might find branch points, in which now divergent organisms converge onto a common ancestor. Likewise with the evolution of a field of inquiry. Since this is a "sourcebook", rather than a straight history, we are guided by our present day theories and approaches, and so are more concerned with *tracing back* various ancestors. Some of these ancestors look like the present day approaches, and others don't. But regardless of which is the case, they have nonetheless been involved in the development of the present approaches. Hence, we have often erred on the side of being too liberal where ideas that originated in a slightly different context were nonetheless incorporated into quantum gravity research at some later date.

Quantum gravity is, of course, yet to be articulated in any final, agreed upon formulation. As alluded to above, with new developments in physics, the quantum gravity project would attempt to avail itself of some potentially relevant feature—it is, thus, a "parasitic" enterprise for much of its early history: wave mechanics, spin and the Dirac equation, neutrinos, the discovery of new forces, mesons and cosmic rays, … . All of these and more were immediately taken up as of potential relevance in quantum gravity's definition and domain. Parts I and II study the ways in which quantum gravity was studied in its embryonic and infancy stages. In cases where it is not parasitic, it is viewed not so much as a problem in its own right, but as an interesting case study, or else a *resource* to cure problems in field theory more generally (the "more serious" business of physics). For example, in 1938 one can find Born writing that there "seems to be a general conviction that the difficulties of our present theory of ultimate particles and nuclear phenomena (the infinite values of the self energy, the zero energy and other quantities) are connected with the problem of merging quantum theory and [general] relativity into a consistent unit".[5] Parts III and IV cover such aspects. Part III also deals more generally with the direct quan-

[4] In the historical literature of general relativity it is known as "the renaissance of GR"—see, e.g., Jean Eisenstaedt's "The Low Water Mark of General Relativity, 1925–1955" (in D. Howard and J. Stachel, eds., *Einstein and the History of General Relativity*, Birkhäuser, 1989: 277–292), David Kaiser, "A ψ is just a ψ? Pedagogy, Practice, and the Reconstitution of General Relativity, 1942-1975" (*Studies in History and Philosophy of Modern Physics* **29**, 1998: 321–338), and AB, Roberto Lalli, and Jürgen Renn, "The Reinvention of General Relativity: A Historiographical Framework for Assessing One Hundred Years of Curved Space-time" (*Isis* **106**, 2015: 598–620).

[5] "A Suggestion for Unifying Quantum Theory and Relativity." *Proceedings of the Royal Society London A* **165**(921), 1938: p. 291. By this stage, Born was thinking of the problem as involving the mixing of quantum principles with the principle of general covariance, rather than general relativity more broadly conceived—this on account of the fact that "gravitation by its order of magnitude is a molar effect and applies only to

tization of general relativity using techniques that had been applied to the electromagnetic field.

What is interesting about these early papers, in terms of the "shut up and calculate" narrative that is often told about physics after the war, is that the papers very rarely step into conceptual waters. One can find none of the preoccupation with the status of observables, the existence of space and time, the meaning of diffeomorphism symmetry, and other such foundational problems that spring up in the 1950s—almost immediately where the papers in this volume stop, in fact. To a certain extent this later development had to do with physicists stepping back and considering the classical theory of general relativity more carefully *from a physical point of view* than had been done previously. The reasons for this are clear: the standard techniques faced technical problems of their own. There was a dawning recognition, towards the end of the first half of the twentieth century, that gravity was simply not like other forces.[6] This recognition brought with it the idea that the problem of quantum gravity will most likely not be resolved through a purely *technical* solution.

masses in bulk, not to the ultimate particles." However, the idea that general relativity might be employed as part of the basic framework of a future theory of elementary particles became popular in later work.

[6]We see the sources of this in Part V in which the general covariance of general relativity was tackled head on in the context of a parameter formalism and the constrained Hamiltonian formulation.

Acknowledgements

This book benefitted greatly from the support of Jürgen Renn, who not only co-organised and funded (via the MPIWG) two highly successful workshops on the early history of quantum gravity (at Pasadena in 2010 and Berlin in 2011), but also helped fund several research projects closely related to this book. We wish to also acknowledge Diana Kormos-Buchwald and Tilman Sauer (now University of Mainz), of the Einstein Papers Project, who were also responsible for the organisation of the Pasadena workshop. The participants in those workshop sessions deserve thanks for their probing analyses of several of the papers that appear in this volume. We would also like to thank Urte Brauckmann of the MPIWG library for her invaluable help in procuring the reproduction rights to the papers. Dean Rickles would also like to thank the Australian Research Council for its support of his historical research via a Future Fellowship: FT130100466.

Dimensions and Domains

Chapter 1
Quantum Theory Meets Gravitation: First Encounters
Dean Rickles

These earliest papers (taken from 1916–1921) suffer from the fact that quantum theory was still in an embryonic state. Quantum mechanics and quantum field theory would have to wait several years before they were even properly formulated, and so could be compared with general relativity (as subsequent parts of this volume will cover). Still, already in these early papers, questions were being posed concerning the proper way to understand the relationship between quantum theory and general relativity. In addition to areas in which there might be modifications brought about by one theory to the other (conflict), or overlaps of some kind (coexistence), there were also "borders" raised between the theories that persisted throughout future changes in physics in the period before 1950. Hence, although one cannot reasonably call this research "quantum gravity," many elements of this pre-history nonetheless directly inspire that later research and, more importantly, serve to demarcate the domains of the ingredient theories.

What is crucial to note in this period is that the developments in quantum theory and in general relativity are often made by the same people, laying the foundations of both simultaneously. For example, as the field equations of general relativity were entering their ultimate form ("the final stage in which the battle over the field equations is being fought out"[1]), Einstein was corresponding with Arnold Sommerfeld about the potential impact of his new theory on contemporary issues in quantum theory. Sommerfeld appears to have thought that general relativity might be of relevance (imposing new constraints and so on) with regard to spectral physics and the Stark effect (Sommerfeld 2000, 438). Einstein writes back:

> General relativity is unlikely to be able to assist you, because it practically coincides with the more restricted theory of relativity for those problems. [...] [A]ny other theory that corresponds with relativity in the restricted sense can be taken over in the general theory of relativity through simple transformation, without the latter delivering any new criteria. Thus you see that I cannot help you in the least. [Einstein, letter to Sommerfeld, [Berlin] 9 December 1915 (Schulmann et al. 1998, 159)]

However, in his paper "Näherungsweise Integration der Feldgleichungen der Gravitation" (1916, Chapter 2 in this volume), Einstein claims that his newly constructed theory of gravitation would itself need to be modified by quantum theory along the same lines as classical electrodynamics so as to keep gravitationally radiating systems stable. Hence, without a quantized emission of gravitational energy, the system would (eventually) continuously radiate away all its energy and collapse. It is, perhaps, no surprise that Einstein was working in tandem on the quantum theory of emission and absorption of radiation at the same time he wrote this paper on gravitational radiation, and the same terminology and concepts can be found in both contexts.

[1] As described in a letter to Sommerfeld, [Berlin] 9 December 1915 (Schulmann et al. 1998, 159).

The Einstein paper features further, and deeper, analogies between electromagnetism and general relativity, and these analogies continued to play a role in later research on quantum (and classical) gravity.[2] For example, just as an accelerated electrically charged particle will emit electromagnetic waves (at the velocity of light), gravitational waves were a direct prediction of the linear theory and were also transverse and propagated at the speed of light. The linear approach (a first approximation to the full non-linear theory, again analogous to electrodynamics' standard perturbative approximation where the coupling constant is taken to be small) involves decomposing the metric $g_{\mu\nu}$ apart as follows:

$$g_{\mu\nu} = -\delta_{\mu\nu} + \gamma_{\mu\nu} \quad (|\gamma_{\mu\nu}| << 1) \tag{1.1}$$

The small perturbations (representing deviations in the spacetime metric away from flat space[3]) $\gamma_{\mu\nu}$ are defined against a flat, Minkowski spacetime background (a "Galilean space" in which $x_4 = it$). The calculation follows exactly the same route as for the retarded electromagnetic potentials, giving the various analogous properties mentioned above, including the existence of gravitational waves.

This perturbative approach (employing a background metric against which the properties of the perturbations are defined) forms one of the standard methods in later work. As we see in Part III, Rosenfeld uses just such a framework in his early quantization and it informs the work of Fierz and Pauli in unpacking the properties of gravitational quanta (massless particles with two helicity states)—Jacques Solomon raised an early warning about the problems faced when ignoring the full non-linear theory; Dmitri Ivanenko noted that during a discussion at the Turin Congress (in 1956), Pauli expressed skepticism about the generalisation of results from the weak field approximation to the full theory in the case of gravitation (Ivanenko 1956, 355).

Einstein published a second paper on gravitational waves in 1918, correcting some problems in the 1916 paper and drawing attention to some crucial differences between the electromagnetic and gravitational cases (which had led Einstein to a serious error of calculation).[4] The most important of these concerned the differences in moments of the radiation: dipole in the case of electrodynamics (as he had also assumed held for gravity in

[2]Gennady Gorelik (1992) briefly discusses the role of the electrodynamical-gravitational analogy in this phase of Einstein's thinking. He argues, rightly we think, that an overly strong analogy persisted in the decades following, fooling physicists into thinking that the task of constructing a quantum theory of gravity was going to be along the same lines as the quantum treatment of electrodynamics. Much of this was an artefact of the linearised approximation scheme Einstein used in 1916, which eradicates the distinguishing (and very difficult) features of gravitation. Given the tractability of the linear approximation, this was the scheme that was used in most of the early direct quantization attempts.

[3]A "deviation of the continuum from one that is field-free" in Einstein's own words (Gorelik 1992).

[4]The prediction of waves would form one of the central motivations for quantum gravity in the 50s. Readers will no doubt be aware of Einstein's double about-face in 1936, leaving him back where he started! Einstein was briefly fooled into thinking that the existence of gravitational waves was an artefact of the linear theory (true only for weak fields), so that when the non-linear field equations were used instead, the waves were no longer physical or energy-transporting. Together with Nathan Rosen, he had originally argued that (cylindrical) gravitational waves are unreal, in a paper entitled "Are Gravitational Waves Real?" (submitted to *Physical Review*). But following Howard Robertson's referee report pointing out a serious flaw in their paper (initially dismissed as erroneous by Einstein), he switched his position (though publishing in a different venue, the *Journal of the Franklin Institute*, and without acknowledgement of Robertson's helpful critique). For a discussion of this episode, see Kennefick (1998)—the title of Kennefick's book (*Travelling at the Speed of Thought*) refers to Eddington's characterisation of gravitational waves in 1922 (Eddington 1922, 269).

Of course, as a practical prediction the result suffered from the fact that there were no known sources (terthe 1916 paper), but quadrupole in the case of gravitation: "[A] mechanical system which permanently retains spherical symmetry cannot radiate" (Einstein 1918, 23). Einstein reiterates the point regarding the likely modifications of general relativity brought about by a completed quantum theory, since the results still indicated energy loss due to thermal agitation (Einstein 1918, 23).

Just one year later Einstein can be found speculating about the role of gravitational fields in the constitution of elementary particles of matter (electrons) in his paper "Do Gravitational Fields Play an Important Role in the Constitution of the Elementary Particles?".[5] The idea is to get discrete particles (electrons) out of a continuous field theory. John Wheeler would later resurrect something similar in the early 1950s with his "Geon project" (where a geon is a "gravitational electromagnetic entity"), in which the elementary particles are reduced to geometrical and topological aspects of such entities (see, e.g. Wheeler (1955)).[6] Once this shift in Einstein's thinking had occurred, he didn't waver again and sought only classical theories as opposed to quantum gravity theories. However, there were additional important developments in Einstein's work. Not least among these was the result, obtained with Infeld and Hoffmann (Einstein, Infeld, and Hoffmann 1938), that enabled particle trajectories (equations of motion) to be derived from the field equations—this would have a direct impact on Peter Bergmann's earliest work on quantum gravity (Bergmann 1949). Bergmann was concerned with using this feature of generally covariant theories to eliminate the divergences caused by particle interactions (infinite interaction terms) in field theories: field singularities (particles) can be determined from the field equations. In this project, it is clear that Bergmann has Einstein's old question, of the constitution of elementary particles, in mind.

The general idea that quantum and gravity would have to meet in some way in a complete formulation of the world seems to have been accepted by other physicists of the time, and motivated subsequent early work and commentaries on the subject. While there were many remarks pointing to the clash, there wasn't much real detailed work to resolve the problem, or even any systematic investigation, until Léon Rosenfeld's efforts in 1930 (following on from initial work on quantum electrodynamics), discussed in the next parts. As mentioned above, until quantum theory was established on a firmer footing, it would have been impossible to tackle the problems of unification.

While the majority of the Eddington paper (Chapter 3), from 1918, amounts to a popular presentation of general relativity[7], it also contains the first statement that we know of concerning the relevance of the Planck length in connection with (quantum) gravita-

---

restrial or astrophysical) that could generate energies of sufficient magnitude to be detected—this changed in the 1960s.

[5]According to Vizgin (2011, 163), this marks the birth of Einstein's switch to "unified field theory," in which gravitational (and electromagnetic) fields are fundamental, and discrete atomic particles emergent features. The basic question of whether and to what extent the gravitational field might be involved in the structure of the electron also plays a part in Paper 3, by Jeffery (see below).

[6]In order to recover such things as charge and spin, the geon project became rather complex, with e.g. wormholes (multiply connected space) being used to recover charge phenomenology. For more on the history, see Dieter Brill (John Wheeler's student): http://mediathek.mpiwg-berlin.mpg.de/mediathekPublic/versionEins/Conferences-Workshops/Quantum-Gravity/Thursday/J-A-Wheeler-Geons.html. The geon was originally called "kugelblitz" (ball lightning) by Wheeler. It is interesting to speculate on the origins of the concept in his work on plasma physics for Project Matterhorn: one of his geon models was a *toroidal* geon of the kind that is often invoked in ball lightning research (where it is modeled as a stable, spinning plasma toroid).

[7]This underplays other important aspects of the paper. For example, it includes a particular perspective on the physical content of the general theory of relativity involving the idea that it is a theory of *space* considered as "the scaffolding constructed from our measures"—a view that retains a degree of modernity.

tional physics, and includes a statement of its importance in any future theories that wish to merge gravitation (cG) and quantum theory (h).[8] Eddington sees that the minute nature of the fundamental length generated from the theoretical constants meant that the structure of a unifying theory will be hidden until we can probe down to the "quadrillionth or quintillionth of a centimetre".[9] Though the numerical value he cites is not as we write it today (perhaps a computational error?), given that he speaks of the fundamental unit length one achieves by combining the three basic universal constants cGh, and there is only one way to do this, we must assume that he has in mind the Planck length nonetheless.

Hence, by 1918 we already have two arguments (however coarsely and briefly expressed) for the necessity of a unification of quantum theory and general relativity (or gravitation and quantum phenomena), both for consistency and for reasons of theoretical unity. It is not yet quantum gravity in the sense of quantization that is being proposed. Indeed, *nothing* is being proposed at this stage; rather, it is left as a future project. Both Einstein and Eddington would soon diverge rather radically from the three motivations outlined earlier. Though Eddington stuck to the project of unification of quantum theory and general relativity, until his death, his approach departed from the reductionist (deeper probing) method he suggests in Paper 2. Perhaps forced by the sheer distance of scales (from known physics) we see in the Planck units, Eddington began to employ a non-experimental methodology, culminating in his (posthumously published) *Fundamental Theory* (Cambridge University Press, 1946).

George Barker Jeffery (9 May 1891—27 April 1957) was one of the few physicists to directly acknowledge Einstein's remarks about the probable modifications that quantum theory would bring about in the theory of gravitational radiation.[10] The main result of the paper (Chapter 4) is not entirely novel, though he appears to have discovered them independently: Jeffery himself notes that he had been unaware until publication that Nordström had only just derived similar results to his. The central outcome is an extension of Schwarzschild's line element (of space surrounding a point) to the combined electromagnetic and gravitational case (considering particles with charge and inertial mass), done by modifying the γ-term as:

$$\gamma = 1 - \frac{2\kappa m}{c^2 r} + \frac{\kappa \epsilon^2}{4\pi c^4 r^2} \tag{1.2}$$

In this case the point is a singularity of both fields (i.e. there are two singularities), rather than the gravitation field alone:

$$ds^2 = -\gamma^{-1} dr^2 - r^2 d\theta^2 - r^2 sin^2\theta d\phi^2 + \gamma c^2 dt^2. \tag{1.3}$$

[8]Hence, here we disagree with Gorelik who assigns priority to Matvey Bronstein in 1936 (see Gorelik (1992, 367)). However, it is true that the first *detailed* analysis of the extent to which the Planck values encode information about the quantum limits of general relativity (in the sense of measurability analysis) was due to Bronstein. Note that the issue of *domains* of applicability of theories would return again, firstly in the context of the debate about measurability limitations (alluded to already), and also in the context of a minimum length implied by quantum theory which are also linked to the former in some cases.

[9]Recall that in England a quintillionth is 10^{-30}. For two more decades, the smallest length relevant to physics was set to 10^{-13} in line with the scales set by the existence of known particles (rather than any future combination of relativity and the quantum).

[10]Like Eddington, Jeffery was a Quaker, and had spent time in prison in 1916 as a conscientious objector—Einstein was aware of this writing wittily in his letter to Jeffery that it "is a highly welcome fact that a considerable portion of England's learned world upholds the pacifist ideal" (Kormos-Buchwald et al. 2009, 85).

Oliver Lodge discusses Jeffery's proposal as an attempt to "ascertain something about the state of the aether close to an electron" (Lodge 1921, 392), which indicates something about the transitionary state of physics at the time. But he also rightly points out that the paper involves thinking about the status of fields near to points, and the question of whether the elementary particles really have structure, which itself leads to the additional question of radiation's interaction with elementary particles in such cases (especially as regards absorption of radiation by elementary particles). As he puts it: "a study of what happens to radiation when it impinges on, or penetrates between the ultimate elements of matter—in fact, a study of the whole behaviour of a stream of radiation at its two ends, the source and the sink—is obviously of great importance" (Lodge 1921, 392). Such reasoning is hardly shot down by the act of bringing those "ultimate elements of matter" in line with quantum principles.

The ultimate aim of Jeffery's paper was to show how the gravitational field might be involved in the structure of the electron (in line with Einstein's "turning point" mentioned above), with the conclusion that the electrical and mass potentials would offer some kind of stabilizing effect by opposing each other.[11] This brings him to Einstein's rather pessimistic remarks about the fate of general relativity at the hands of quantum theory from the 1916 paper (Chapter 2), regarding the instability of the atom in the face of continuous classical gravitational radiation. Jeffery believed that the problem could be evaded. Einstein was not convinced by Jeffery's idea:

> I unfortunately cannot share your optimism regarding the solution to the quantum problem. I believe that the theory of relativity does not bring us a step closer, at least in its current form. I am convinced that the two-body problem will not lead to a discrete manifold of paths but to a continuous one. (Einstein, letter to Jeffery [Berlin], 18 March 1921 (Kormos-Buchwald et al. 2009, 85))

But, as we have seen, Einstein had already started down a path that followed the spirit if not the letter of Jeffery's approach (namely, using GR, or some modification of it, to recover quantum behaviour). In other words, Einstein's earlier assumption about quantum restrictions of gravitation were replaced by the view that quantum phenomena are to be derived from general relativity and thus are not fundamental. However, here Einstein does not indicate this aspect of his thinking to Jeffery, and indeed the quote above looks largely negative as far as the entire project of getting quantum from relativity goes.

Though Einstein was followed by several others along the "unified field theory" path, the majority view was that his earlier pessimism should and could be responded to without rejecting quantum theory, as we see in the papers that follow. In the next part, the two

[11] This, he thought, might be the result of an analysis of the two body problem in his scheme: a subject for future work. Of course, this problem (exact solutions for the gravitational two-body problem for two point singularities) would ultimately have to wait for new methods in the initial value problem and numerical techniques—though the Einstein, Hoffmann, Infeld paper (1938) is precisely along these lines, and, as we have seen, Bergmann used this method precisely for its ability to avoid interaction divergences. Curiously, Jeffery argues that without including electric charge, the point singularity blows up so that one has "not a solution with a single point singularity, but a solution with a point singularity surrounded by a spherical surface of singularity" (see p. 42). Adding charge eliminates all zeros and infinities other than that occurring at $r = 0$. Even more curiously, Jeffery speculates that given the light deflection features of general relativity "it would seem to be not impossible that a ray which passed sufficiently close to an attracting particle might be so strongly deflected that it would be permanently entrapped by the particle" (see p. 42). Again, the introduction of charge is invoked to prevent such light sinks (it seems clear that something like a black hole is being suggested here).

ingredient theories are worked out in more detail along exactly the lines of figuring out *how* generally relativistic principles and quantum principles can co-exist, at least in a formal-structural sense.

References

Bergmann, Peter G. (1949). Non-Linear Field Theories. *Physical Review* 75(4):680–685.

Eddington, Arthur S. (1922). The Propagation of Gravitation Waves. *Proceedings of the Royal Society of London A* 102(716):268–282.

Einstein, Albert (1918). Über Gravitationswellen. *Sitzungsberichte der Königlich Preußischen Akademie der Wissenschaften Berlin*:154–167.

Einstein, Albert, Leopold Infeld, and Banesh Hoffmann (Jan. 1938). The Gravitational Equations and the Problem of Motion. *Annals of Mathematics* 39(1):65–100.

Gorelik, Gennady E. (1992). First Steps of Quantum Gravity and the Planck Values. In: *Studies in the History of General Relativity*. Ed. by Jean Eisenstaedt and A. J. Kox. Boston: Birkhäuser.

Ivanenko, Dmitri (1956). Non-Linear Generalisations of the Field Theory and the Constant of Minimal Length. *Supplemento al Nuovo Cimento* 6:349–355.

Kennefick, Daniel (1998). History of the Radiation Reaction Problem. In: *The Expanding Worlds of General Relativity*. Ed. by Hubert Goenner, Jürgen Renn, Jim Ritter, and Tilman Sauer. Boston: Birkhäuser.

Kormos-Buchwald, Diana L., Ze'ev Rosenkranz, Tilman Sauer, József Illy, and Virginia I. Holmes (2009). *The Collected Papers of Albert Einstein,* Vol. 12. Princeton: Princeton University Press.

Lodge, Oliver (1921). The Gravitational Field of an Electron. *Nature* 2691(107):392.

Schulmann, Robert, A.J. Kox, Michel Janssen, and József Illy (1998). *The Collected Papers of Albert Einstein, Volume 8*. Princeton: Princeton University Press.

Sommerfeld, Arnold (2000). *Wissenschaftlicher Briefwechsel. Band 1: 1892–1918*. München: GNT-Verlag.

Vizgin, Vladimir P. (2011). *Unified Field Theories: In the First Third of the 20th Century (translated by Julian Barbour)*. Boston: Birkhäuser.

Wheeler, John A. (1955). Geons. *Physical Review* 97(2):511–536.

Chapter 2
Albert Einstein (1916): Näherungsweise Integration der Feldgleichungen der Gravitation

Albert Einstein (1916). Näherungsweise Integration der Feldgleichungen der Gravitation. *Königlich Preußische Akademie der Wissenschaften* (Berlin). *Sitzungsberichte*.

Näherungsweise Integration der Feldgleichungen der Gravitation.

Von A. Einstein.

Bei der Behandlung der meisten speziellen (nicht prinzipiellen) Probleme auf dem Gebiete der Gravitationstheorie kann man sich damit begnügen, die $g_{\mu\nu}$ in erster Näherung zu berechnen. Dabei bedient man sich mit Vorteil der imaginären Zeitvariable $x_4 = it$ aus denselben Gründen wie in der speziellen Relativitätstheorie. Unter »erster Näherung« ist dabei verstanden, daß die durch die Gleichung

$$g_{\mu\nu} = -\delta_{\mu\nu} + \gamma_{\mu\nu} \tag{1}$$

definierten Größen $\gamma_{\mu\nu}$, welche linearen orthogonalen Transformationen gegenüber Tensorcharakter besitzen, gegen 1 als kleine Größen behandelt werden können, deren Quadrate und Produkte gegen die ersten Potenzen vernachlässigt werden dürfen. Dabei ist $\delta_{\mu\nu} = 1$ bzw. $\delta_{\mu\nu} = 0$, je nachdem $\mu = \nu$ oder $\mu \neq \nu$.

Wir werden zeigen, daß diese $\gamma_{\mu\nu}$ in analoger Weise berechnet werden können wie die retardierten Potentiale der Elektrodynamik. Daraus folgt dann zunächst, daß sich die Gravitationsfelder mit Lichtgeschwindigkeit ausbreiten. Wir werden im Anschluß an diese allgemeine Lösung die Gravitationswellen und deren Entstehungsweise untersuchen. Es hat sich gezeigt, daß die von mir vorgeschlagene Wahl des Bezugssystems gemäß der Bedingung $g = |g_{\mu\nu}| = -1$ für die Berechnung der Felder in erster Näherung nicht vorteilhaft ist. Ich wurde hierauf aufmerksam durch eine briefliche Mitteilung des Astronomen de Sitter, der fand, daß man durch eine andere Wahl des Bezugssystems zu einem einfacheren Ausdruck des Gravitationsfeldes eines ruhenden Massenpunktes gelangen kann, als ich ihn früher gegeben hatte[1]. Ich stütze mich daher im folgenden auf die allgemein invarianten Feldgleichungen.

[1] Sitzungsber. XLVII, 1915, S. 833.

§ 1. Integration der Näherungsgleichungen des Gravitationsfeldes.

Die Feldgleichungen lauten in ihrer kovarianten Form

$$\left.\begin{aligned} R_{\mu\nu} + S_{\mu\nu} &= -\varkappa\left(T_{\mu\nu} - \frac{1}{2} g_{\mu\nu} T\right) \\ R_{\mu\nu} &= -\sum_\alpha \frac{\partial}{\partial x_\alpha}\begin{Bmatrix}\mu\nu \\ \alpha\end{Bmatrix} + \sum_{\alpha\beta}\begin{Bmatrix}\mu\alpha \\ \beta\end{Bmatrix}\begin{Bmatrix}\nu\beta \\ \alpha\end{Bmatrix} \\ S_{\mu\nu} &= \frac{\partial \log\sqrt{g}}{\partial x_\mu \partial x_\nu} \sum_\alpha \begin{Bmatrix}\mu\nu \\ \alpha\end{Bmatrix} \frac{\partial \log\sqrt{g}}{\partial x_\alpha} \end{aligned}\right\} \quad (1)$$

Dabei bedeuten die geschweiften Klammern die bekannten CHRISTOFFELschen Symbole, $T_{\mu\nu}$ den kovarianten Energietensor der Materie, T den zugehörigen Skalar. Die Gleichungen (1) liefern in der uns interessierenden Näherung die durch Entwickeln unmittelbar folgenden Gleichungen

$$\sum_\alpha \frac{\partial^2 \gamma_{\mu\alpha}}{\partial x_\nu \partial x_\alpha} + \sum_\alpha \frac{\partial^2 \gamma_{\nu\alpha}}{\partial x_\mu \partial x_\alpha} - \sum_\alpha \frac{\partial^2 \gamma_{\mu\nu}}{\partial x_\alpha^2} - \frac{\partial^2}{\partial x_\mu \partial x_\nu}\left(\sum_\alpha \gamma_{\alpha\alpha}\right) = -2\varkappa\left(T_{\mu\nu} - \frac{1}{2}\delta_{\mu\nu}\sum T_{\alpha\alpha}\right). \quad (2)$$

Das letzte Glied der linken Seite stammt von der Größe $S_{\mu\nu}$, die bei der von mir bevorzugten Koordinatenwahl verschwindet. Die Gleichungen (2) lassen sich durch den Ansatz

$$\gamma_{\mu\nu} = \gamma'_{\mu\nu} + \psi\delta_{\mu\nu} \quad (3)$$

lösen, wobei die $\gamma'_{\mu\nu}$ der zusätzlichen Bedingung

$$\sum_\nu \frac{\partial \gamma'_{\mu\nu}}{\partial x_\nu} = 0 \quad (4)$$

genügen. Durch Einsetzen von (3) in (2) erhält man an Stelle der linken Seite

$$-\sum_\alpha \frac{\partial \gamma'_{\mu\nu}}{\partial x_\alpha^2} - \frac{\partial^2}{\partial x_\mu \partial x_\nu}\left(\sum_\alpha \gamma'_{\alpha\alpha}\right) + 2\frac{\partial^2\psi}{\partial x_\mu \partial x_\nu} - \delta_{\mu\nu}\sum_\alpha \frac{\partial^2\psi}{\partial x_\alpha^2} - 4\frac{\partial^2\psi}{\partial x_\mu \partial x_\nu}.$$

Der Beitrag des zweiten, dritten und fünften Gliedes verschwindet, wenn ψ gemäß der Gleichung

$$\sum_\alpha \gamma'_{\alpha\alpha} + 2\psi = 0 \quad (5)$$

gewählt wird, was wir festsetzen. Mit Rücksicht hierauf erhält man an Stelle von (2)

$$\sum_\alpha \frac{\partial^2}{\partial x_\alpha^2}\left(\gamma'_{\mu\nu} - \frac{1}{2}\delta_{\mu\nu}\sum_\alpha \gamma'_{\alpha\alpha}\right) = 2\varkappa\left(T_{\mu\nu} - \frac{1}{2}\delta_{\mu\nu}\sum_\alpha T_{\alpha\alpha}\right)$$

oder

$$\sum_\alpha \frac{\partial^2}{\partial x_\alpha^2}\gamma'_{\mu\nu} = 2\varkappa T_{\mu\nu}. \tag{6}$$

Es ist hierzu zu bemerken, daß Gleichung (6) mit der Gleichung (4) im Einklang ist. Denn es ist zunächst leicht zu zeigen, daß bei der von uns erstrebten Genauigkeit der Impulsenergiesatz für die Materie durch die Gleichung

$$\sum_\nu \frac{\partial T_{\mu\nu}}{\partial x_\nu} = 0 \tag{7}$$

ausgedrückt ward. Führt man an (6) die Operation $\sum_\nu \frac{\partial}{\partial x_\nu}$ aus, so verschwindet nicht nur vermöge (4) die linke Seite, sondern, wie es sein muß, vermöge (7) auch die rechte Seite von (6). Wir merken an, daß wegen (3) und (5) die Gleichungen

$$\gamma_{\mu\nu} = \gamma'_{\mu\nu} - \frac{1}{2}\delta_{\mu\nu}\sum_\alpha \gamma'_{\alpha\alpha} \tag{8}$$

$$\gamma'_{\mu\nu} = \gamma_{\mu\nu} - \frac{1}{2}\delta_{\mu\nu}\sum_\alpha \gamma_{\alpha\alpha} \tag{8a}$$

bestehen. Da sich die $\gamma'_{\mu\nu}$ nach Art der retardierten Potentiale berechnen lassen, so ist damit unsere Aufgabe gelöst. Es ist

$$\gamma'_{\nu\mu} = -\frac{\varkappa}{2\pi}\int \frac{T_{\mu\nu}(x_0, y_0, z_0, t-r)}{r}\, dV_0. \tag{9}$$

Dabei sind mit x, y, z, t die reellen Koordinaten $x_1, x_2, x_3, \frac{x_4}{i}$ bezeichnet, und zwar bezeichnen sie ohne Indizes die Koordinaten des Aufpunktes, mit dem Index »o« diejenigen des Integrationselementes. dV_0 ist das dreidimensionale Volumelement des Integrationsraumes r der räumliche Abstand $\sqrt{(x-x_0)^2+(y-y_0)^2+(z-z)^2}$.

Für das Folgende bedürfen wir ferner der Energiekomponenten des Gravitationsfeldes. Wir erhalten sie am einfachsten direkt aus den Gleichungen (6). Durch Multiplikation mit $\frac{\partial \gamma_{\mu\nu}}{\partial x_\sigma}$ und Summation über μ und ν erhält man auf der linken Seite nach geläufiger Umformung

$$\frac{\partial}{\partial x_\alpha}\left[\sum_{\mu\nu}\frac{\partial \gamma'_{\mu\nu}}{\partial x_\sigma}\frac{\partial \gamma'_{\mu\nu}}{\partial x_\alpha} - \frac{1}{2}\partial_{\sigma\alpha}\sum_{\mu\nu\beta}\left(\frac{\partial \gamma'_{\mu\nu}}{\partial x_\beta}\right)^2\right].$$

Diese Klammergröße drückt bis auf den Proportionalitätsfaktor offenbar die Energiekomponenten $t_{\sigma\alpha}$ aus; der Faktor ergibt sich leicht durch Berechnen der rechten Seite. Der Impuls-Energie-Satz der Materie lautet ohne Vernachlässigungen

$$\sum_{\sigma} \frac{\partial \sqrt{-g}\, T_\mu^\sigma}{\partial x_\sigma} + \frac{1}{2} \sum_{\varrho\sigma} \frac{\partial g^{\varrho\sigma}}{\partial x_\mu} \sqrt{-g}\, T_{\varrho\sigma} = 0 .$$

Mit dem von uns gewünschten Grade der Näherung kann man dafür setzen

$$\sum_{\sigma} \frac{\partial T_{\mu\sigma}}{\partial x_\sigma} + \frac{1}{2} \sum_{\varrho\sigma} \frac{\partial g_{\varrho\sigma}}{\partial x_\mu} T_{\varrho\sigma} = 0 . \qquad (7\text{a})$$

Es ist dies die um einen Grad exaktere Formulierung zu Gleichung (7). Hieraus folgt, daß die rechte Seite von (6) bei der ins Auge gefaßten Umformung

$$-4\varkappa \sum \frac{\partial T_{\mu\nu}}{\partial x_\nu}$$

liefert. Der Erhaltungssatz lautet also

$$\sum \frac{\partial (T_{\mu\nu} + t_{\mu\nu})}{\partial x_\nu} = 0 , \qquad (10)$$

wobei

$$t_{\mu\nu} = \frac{1}{4\varkappa} \left[\sum_{\alpha\beta} \frac{\partial \gamma'_{\alpha\beta}}{\partial x_\mu} \frac{\partial \gamma'_{\alpha\beta}}{\partial x_\nu} - \frac{1}{2} \delta_{\mu\nu} \sum_{\alpha\beta\tau} \left(\frac{\partial \gamma'_{\alpha\beta}}{\partial x_\tau} \right)^2 \right] \qquad (11)$$

die Energiekomponenten des Gravitationsfeldes sind.

Als einfachstes Anwendungsbeispiel berechnen wir das Gravitationsfeld eines im Koordinatenursprung ruhenden Massenpunktes von der Masse M. Der Energietensor der Materie ist bei Vernachlässigung der Flächenkräfte durch

$$T_{\mu\nu} = \rho \frac{dx_\mu}{ds} \frac{dx_\nu}{ds} \qquad (12)$$

gegeben, mit Rücksicht darauf, daß in erster Näherung der kovariante Tensor der Energie durch den kontravarianten ersetzt werden kann. Der Skalar ρ ist die (natürlich gemessene) Massendichte. Es ergibt sich aus (9) und (12), daß alle $\gamma'_{\mu\nu}$ bis auf γ'_{44} verschwinden, für welch letztere Komponente sich ergibt

$$\gamma'_{44} = -\frac{\varkappa}{2\pi} \frac{M}{r} . \qquad (13)$$

Hieraus erhält man mit Hilfe von (8) und (1) für die $g_{\mu\nu}$ die Werte

$$\left.\begin{matrix} -1-\frac{\varkappa}{4\pi}\frac{M}{r} & 0 & 0 & 0 \\ 0 & -1-\frac{\varkappa}{4\pi}\frac{M}{r} & 0 & 0 \\ 0 & 0 & -1-\frac{\varkappa}{4\pi}\frac{M}{r} & 0 \\ 0 & 0 & 0 & -1-\frac{\varkappa}{4\pi}\frac{M}{r} \end{matrix}\right\} \qquad (14)$$

Diese Werte, welche sich von den von mir früher angegebenen nur vermöge der Wahl des Bezugssystems unterscheiden, wurden mir durch Hrn. de Sitter brieflich mitgeteilt. Sie führten mich auf die im vorstehenden angegebene einfache Näherungslösung. Es ist aber wohl im Auge zu behalten, daß der hier benutzten Koordinatenwahl keine entsprechende im allgemeinen Falle zur Seite steht, indem die $\gamma_{\mu\nu}$ und $\gamma'_{\mu\nu}$ nicht beliebigen, sondern nur linearen, orthogonalen Substitutionen gegenüber Tensorcharakter besitzen.

§ 2. Ebene Gravitationswellen.

Aus den Gleichungen (6) und (9) folgt, daß sich Gravitationsfelder stets mit der Geschwindigkeit 1, d. h. mit Lichtgeschwindigkeit, fortpflanzen. Ebene, nach der positiven x-Achse fortschreitende Gravitationswellen sind daher durch den Ansatz zu finden

$$\gamma'_{\mu\nu} = \alpha_{\mu\nu} f(x_1 + i x_4) = \alpha_{\mu\nu} f(x-t) . \qquad (15)$$

Dabei sind die $\alpha_{\mu\nu}$ Konstante; f ist eine Funktion des Arguments $x-t$. Ist der betrachtete Raum frei von Materie, d. h. verschwinden die $T_{\mu\nu}$, so sind die Gleichungen (6) durch diesen Ansatz erfüllt. Die Gleichungen (4) liefern zwischen den $\alpha_{\mu\nu}$ die Beziehungen

$$\left.\begin{aligned} \alpha_{11}+i\alpha_{14} &= 0 \\ \alpha_{12}+i\alpha_{24} &= 0 \\ \alpha_{13}+i\alpha_{34} &= 0 \\ \alpha_{14}+i\alpha_{44} &= 0 \end{aligned}\right\} . \qquad (16)$$

Von den 10 Konstanten $\alpha_{\mu\nu}$ sind daher nur 6 frei wählbar. Wir können die allgemeinste Welle der betrachteten Art daher aus Wellen von folgenden 6 Typen superponieren

$$\left.\begin{matrix} \text{a)}\ \begin{matrix}\alpha_{11}+i\alpha_{14}=0 \\ \alpha_{14}+i\alpha_{44}=0\end{matrix} & \begin{matrix}\text{b)}\ \alpha_{12}+i\alpha_{24}=0 \\ \text{c)}\ \alpha_{13}+i\alpha_{34}=0\end{matrix} & \begin{matrix}\text{d)}\ \alpha_{22}\neq 0 \\ \text{e)}\ \alpha_{23}\neq 0 \\ \text{f)}\ \alpha_{33}\neq 0\end{matrix} \end{matrix}\right\} . \qquad (17)$$

Diese Angaben sind so aufzufassen, daß für jeden Typ die in seinen Bedingungen nicht explizite genannten $\alpha_{\mu\nu}$ verschwinden; im Typ a sind also nur α_{11}, α_{14}, α_{44} von null verschieden usw. Den Symmetrieeigenschaften nach entspricht Typ a einer Longitudinalwelle, die Typen b und c Transversalwellen, während die Typen d, e, f einem neuartigen Symmetriecharakter entsprechen. Die Typen b und c unterscheiden sich nicht im Wesen, sondern nur durch ihre Orientierung gegen die y- und z-Achse voneinander, ebenso die Typen d, e, f, so daß eigentlich drei wesentlich verschiedene Wellentypen existieren.

Uns interessiert in erster Linie die von diesen Wellen transportierte Energie, welche durch den Energiestrom $\mathfrak{f}_x = \frac{1}{i} t_{41}$ gemessen wird. Es ergibt sich aus (11) für die einzelnen Typen.

$$\text{a)}\quad \frac{1}{i} t_{41} = \frac{f'^2}{4\varkappa}(\alpha_{11}^2 + \alpha_{14}^2 + \alpha_{41}^2 + \alpha_{44}^2) = 0$$

$$\text{b)}\quad \frac{1}{i} t_{41} = \frac{f'^2}{4\varkappa}(\alpha_{12}^2 + \alpha_{24}^2) = 0$$

$$\text{c)}\quad \frac{1}{i} t_{41} = \frac{f'^2}{4\varkappa}(\alpha_{13}^2 + \alpha_{34}^2) = 0$$

$$\text{d)}\quad \frac{1}{i} t_{22} = \frac{f'^2}{4\varkappa}\alpha_{22}^2 = \frac{1}{4\varkappa}\left(\frac{\partial \gamma'_{22}}{\partial t}\right)^2$$

$$\text{e)}\quad \frac{1}{i} t_{23} = \frac{f'^2}{4\varkappa}\alpha_{23}^2 = \frac{1}{4\varkappa}\left(\frac{\partial \gamma'_{23}}{\partial t}\right)^2$$

$$\text{f)}\quad \frac{1}{i} t_{33} = \frac{f'^2}{4\varkappa}\alpha_{33}^2 = \frac{1}{4\varkappa}\left(\frac{\partial \gamma'_{33}}{\partial t}\right)^2$$

Es ergibt sich also, daß nur die Wellen des letzten Typs Energie transportieren, und zwar ist der Energietransport einer beliebigen ebenen Welle gegeben durch

$$\mathfrak{l}_x = \frac{1}{i} t_{41} = \frac{1}{4\varkappa}\left[\left(\frac{\partial \gamma'_{22}}{\partial t}\right)^2 + 2\left(\frac{\partial \gamma'_{23}}{\partial t}\right)^2 + \left(\frac{\partial \gamma'_{33}}{\partial t}\right)^2\right]. \qquad (18)$$

§ 3. Energieverlust körperlicher Systeme durch Emission von Gravitationswellen.

Das System, dessen Ausstrahlung untersucht werden soll, befinde sich dauernd in der Umgebung des Koordinatenursprungs. Wir betrachten das vom System erzeugte Gravitationsfeld lediglich für Aufpunkte, deren Abstand R vom Koordinatenursprung groß ist gegenüber den Abmessungen des Systems. Der Aufpunkt werde in die positive x-Achse verlegt, d. h. es sei

$$x_1 = R\,, \quad x_2 = x_3 = 0\,.$$

Die Frage ist dann, ob im Aufpunkt eine nach der positiven x-Achse gerichtete Wellenstrahlung vorhanden ist, welche Energie transportiert. Die Betrachtungen des § 2 zeigen, daß eine solche Strahlung im Aufpunkt nur den Komponenten γ'_{22}, γ'_{23}, γ'_{33} geliefert werden kann. Diese allein haben wir also zu berechnen. Aus (9) ergibt sich

$$\gamma'_{22} = -\frac{\varkappa}{2\pi}\int \frac{T_{22}(x_0, y_0, z_0, t-r)}{r}\, dV_0 .$$

Ist das System wenig ausgedehnt und sind seine Energiekomponenten nicht allzu rasch veränderlich, so kann ohne merklichen Fehler das Argument $t-r$ durch das bei der Integration konstante $t-R$ ersetzt werden. Ersetzt man außerdem $\frac{1}{r}$ durch $\frac{1}{R}$, so erhält man die in den meisten Fällen genügende Näherungsgleichung

$$\gamma'_{22} = -\frac{\varkappa}{2\pi R}\int T_{22}\, dV_0 , \qquad (19)$$

wobei die Integration in gewöhnlicher Weise, d. h. bei konstantem Zeitargument zu nehmen ist. Dieser Ausdruck läßt sich vermittels (7) durch einen für die Berechnung bei materiellen Systemen bequemeren ersetzen. Aus

$$\frac{\partial T_{21}}{\partial x_1} + \frac{\partial T_{22}}{\partial x_2} + \frac{\partial T_{23}}{\partial x_3} + \frac{\partial T_{24}}{\partial x_4} = 0$$

folgt durch Multiplikation mit x_2 und Integration über das ganze System nach partieller Integration des zweiten Gliedes

$$-\int T_{22}\, dV + \frac{\partial}{\partial x_4}\left(\int T_{24} x_2\, dV\right) = 0 . \qquad (20)$$

Ferner folgt aus

$$\frac{\partial T_{41}}{\partial x_1} + \frac{\partial T_{42}}{\partial x_2} + \frac{\partial T_{43}}{\partial x_3} + \frac{\partial T_{44}}{\partial x_4} = 0$$

durch Multiplikation mit $\frac{x_2^2}{2}$ auf analogem Wege

$$-\int T_{24} x_2\, dV + \frac{\partial}{\partial x_4}\left(\int T_{44} \frac{x_2^2}{2}\, dV\right) = 0 . \qquad (21)$$

Aus (20) und (21) folgt

$$\int T_{22}\, dV = \frac{\partial^2}{\partial x_4^2}\left(\int T_{44} \frac{x_2^2}{2}\, dV\right)$$

oder, indem man reelle Koordinaten einführt, und indem man sich die Näherung gestattet, die Energiedichte $(-T_{44})$ auch für beliebig bewegte Massen der ponderabeln Dichte ρ gleichzusetzen

$$\int T_{22} dV = \frac{1}{2} \frac{\partial^2}{\partial t^2} \left(\int \rho y^2 dV \right). \tag{22}$$

Man hat also auch

$$\gamma'_{22} = - \frac{\varkappa}{4\pi R} \frac{\partial^2}{\partial t^2} \left(\int \rho y^2 dV \right). \tag{23}$$

Auf analoge Weise berechnet man

$$\gamma'_{33} = - \frac{\varkappa}{4\pi R} \frac{\partial^2}{\partial t^2} \left(\int \rho z^2 dV \right) \tag{23a}$$

$$\gamma'_{23} = - \frac{\varkappa}{4\pi R} \frac{\partial^2}{\partial t^2} \left(\int \rho y^2 dV \right). \tag{23b}$$

Die in (23), (23a) und (23b) auftretenden Integrale, welche nichts anderes sind als zeitlich variable Trägheitsmomente, nennen wir im folgenden zur Abkürzung J_{22}, J_{33}, J_{23}. Dann ergibt sich für die Intensität $\mathfrak{f}_x$ der Energiestrahlung aus (18)

$$\mathfrak{f}_x = \frac{\varkappa}{64\pi^2 R^2} \left[\left(\frac{\partial^3 J_{22}}{\partial t^3} \right)^2 + 2 \left(\frac{\partial^3 J_{23}}{\partial t^3} \right)^2 + \left(\frac{\partial^3 J_{33}}{\partial t^3} \right)^2 \right]. \tag{20}$$

Hieraus ergibt sich weiter, daß die mittlere Energiestrahlung nach allen Richtungen gegeben ist durch

$$\frac{\varkappa}{64\pi^2 R^2} \cdot \frac{2}{3} \sum_{\alpha\beta} \left(\frac{\partial^3 J_{\alpha\beta}}{\partial t^3} \right)^2,$$

wobei über alle 9 Kombinationen der Indizes 1—3 zu summieren ist. Denn dieser Ausdruck ist einerseits invariant gegenüber räumlichen Drehungen des Koordinatensystems, wie leicht aus dem (dreidimensionalen) Tensorcharakter von $J_{\alpha\beta}$ folgt; anderseits stimmt er im Falle radialer Symmetrie ($J_{11} = J_{22} = J_{33}$; $J_{23} = J_{31} = J_{12} = 0$) mit (20) überein. Man erhält aus ihm also die Ausstrahlung A des Systems pro Zeiteinheit durch Multiplikation mit $4\pi R^2$:

$$A = \frac{\varkappa}{24\pi} \sum_{\alpha\beta} \left(\frac{\partial^3 J_{\alpha\beta}}{\partial t^3} \right)^2. \tag{21}$$

Würde man die Zeit in Sekunden, die Energie in Erg messen, so würde zu diesem Ausdruck der Zahlenfaktor $\frac{1}{c^4}$ hinzutreten. Berücksichtigt man außerdem, daß $\varkappa = 1.87 \cdot 10^{-27}$, so sieht man, daß A in allen nur denkbaren Fällen einen praktisch verschwindenden Wert haben muß.

Gleichwohl müßten die Atome zufolge der inneratomischen Elektronenbewegung nicht nur elektromagnetische, sondern auch Gravitationsenergie ausstrahlen, wenn auch in winzigem Betrage. Da dies in Wahrheit in der Natur nicht zutreffen dürfte, so scheint es, daß die Quantentheorie nicht nur die Maxwellsche Elektrodynamik, sondern auch die neue Gravitationstheorie wird modifizieren müssen.

Nachtrag. Das seltsame Ergebnis, daß Gravitationswellen existieren sollen, welche keine Energie transportieren (Typen a, b, c), klärt sich in einfacher Weise auf. Es handelt sich nämlich dabei nicht um »reale« Wellen, sondern um »scheinbare« Wellen, die darauf beruhen, daß als Bezugssystem ein wellenartig zitterndes Koordinatensystem benutzt wird. Dies sieht man bequem in folgender Weise ein. Wählt man das Koordinatensystem in gewohnter Weise von vornherein so, daß $\sqrt{g} = 1$ ist, so erhält man statt (2) als Feldgleichungen bei Abwesenheit von Materie

$$\sum_\alpha \frac{\partial^2 \gamma_{\mu\alpha}}{\partial x_\nu \partial x_\alpha} + \sum_\alpha \frac{\partial^2 \gamma_{\nu\alpha}}{\partial x_\mu \partial x_\alpha} - \sum_\alpha \frac{\partial^2 \gamma_{\mu\nu}}{\partial x_\alpha^2} = 0 .$$

Führt man in diese Gleichungen direkt den Ansatz

$$\gamma_{\mu\nu} = \alpha_{\mu\nu} f(x_1 + ix_4)$$

ein, so erhält man zwischen den Konstanten $\alpha_{\mu\nu}$ 10 Gleichungen, aus denen hervorgeht, daß nur α_{22}, α_{33} und α_{23} von null verschieden sein können (wobei $\alpha_{22} + \alpha_{33} = 0$). Bei dieser Wahl des Bezugssystems existieren also nur diejenigen Wellentypen (d, e, f), welche Energie transportieren. Die übrigen Wellentypen lassen sich also durch diese Koordinatenwahl wegschaffen; sie sind in dem angegebenen Sinne nicht »wirkliche« Wellen.

Wenn es also auch in dieser Untersuchung sich als bequem herausgestellt hat, die Wahl des Koordinatensystems von vornherein keiner Beschränkung zu unterwerfen, wenn es sich um die Berechnung der ersten Näherung handelt, so zeigt unser letztes Ergebnis doch, daß der Koordinatenwahl gemäß der Bedingung $\sqrt{-g} = 1$ eine tiefe physikalische Berechtigung zukommt.

Ausgegeben am 29. Juni.

Chapter 3
Arthur Stanley Eddington (1918): Gravitation and the Principle of Relativity II

Arthur Stanley Eddington (1918). Gravitation and the Principle of Relativity II. *Nature* 101(2524): 34–36.

professors with fourteen assistant lecturers, and the students in all departments were 556 in 1916. The fees for instruction are 30 francs for each semester, together with certain special fees, and foreigners are charged treble fees. The six schools have a total of 2427 students, and are fully equipped with laboratories for experimental instruction. On leaving these schools the students enter the Union of Swiss Technicians, which association now counts its members by thousands and has for its organ the *Swiss Technical Review*, which publishes much good original work. Altogether these institutions have proved a great success and have been of material benefit in training a large body of men for the industries, many of whom have afterwards qualified for high industrial or administrative positions at home or abroad.

GRAVITATION AND THE PRINCIPLE OF RELATIVITY.[1]

II.

WE have to admit, then, that a world-line can be bent by the proximity of other world-lines. It can also be bent, as you see, by the proximity of my thumb. The suggestion arises, May not the two modes of bending be essentially the same? The bending by my thumb (a mathematical transformation of space and time) is in a sense spurious; the world-line is pursuing a course which is straight relative to the *original* material. Or we may perhaps best put it this way—the world-line still continues to take the shortest path between two points, only it reckons distance according to the length that would be occupied in the unstretched state of the bladder. It is suggested that the deflection of a world-line by gravitation is of the same nature; from each world-line a state of distortion radiates, as if from a badly puckered seam, and any other world-line takes the shortest course through this distorted region, which would immediately become straight if the strain could be undone. The same rule—of shortest distance as measured in the undistorted state—is to hold in all cases. This is a mode of reasoning which has often been fruitful in scientific generalisations. A magnetic needle turns towards the end of a bar-magnet; it also turns towards a spot near the pole of the earth; hence the suggestion that the earth is a magnet. We assume the essential identity of the two modes of deflecting the needle. It is a daring step to apply the analogy and assume the essential identity of the two ways of deflecting world-lines; but at any rate we shall make this assumption and see what comes of it.

You will see that according to this view the earth moves in a curved orbit, not because the sun exerts any direct pull, but because the earth is trying to find the shortest way through a space and time which have been tangled up by an influence radiating from the sun. We can continue to describe this indirect influence of the sun on the earth's motion as a "force"; but, assuming that it makes itself felt as a modification or strain of space and time, we are able to bring the discussion of the laws of this force into line with the discussion of the laws of space and time, *i.e.* the laws of geometry. Needless to say, we could not determine a physical law like the law of gravitation by geometrical reasoning without making some assumption.

I am afraid that to talk of a force as being a distortion of space and time must at first appear to you hopeless jargon. But it must be remembered first that we are not concerned with any metaphysical space and time. We mean by space and time simply a scaffolding that we construct as the result of our measures; and if anything queer happens to our measuring apparatus, the scaffolding may easily go crooked. Taking our everyday conception of space, we should say that this room is at rest; we have been told that it is being carried round the earth once a day, but in practical life we never pay any attention to that. The space that we naturally use is thus different from, and it is not difficult to show that it is distorted as compared with, the more fundamental astronomical space in which this room is travelling at a great velocity. So our scaffolding is crooked. But, it may be asked, in what way can this distortion of our space-scaffolding be regarded as a force? The answer is quite simple. We perceive it as a force, and that is the only way in which we do perceive it. We do not perceive that this room is being carried round by the earth's rotation, but we perceive a certain force—the earth's centrifugal force. It is rather difficult to demonstrate this force, because gravitation predominates overwhelmingly; but if gravity were annihilated we should have to be tied down to the floor to prevent our flying up to the ceiling, and we should certainly feel ourselves pulled by a very vigorous centrifugal force. That is our only perception of the crookedness of our scaffolding.

We often call the centrifugal force an "unreal" force, meaning that it arises simply from a transformation of the framework of reference. Can we feel confident that gravitation is in any sense more "real"? In effect they are so much alike that even in scientific work we speak of them in one breath. What is called the value of gravity in London, 981·17 cm./sec.^{2}, is really made up partly of the true attraction of the earth and partly of the centrifugal force. It is not considered worth while to make any distinction. Surely, then, it is not a great stretch of the imagination to regard gravitation as of the same nature as centrifugal force, being merely our perception of the crookedness of the scaffolding that we have chosen.

If gravity and centrifugal force are manifestations of the same underlying condition, it must be possible to reduce them to the same laws; but we must express the laws in a manner which will render them comparable. There is a convenient form of Newton's law, which was given by Laplace and is well known to mathematicians, which describes how the intensity at any point is related to the intensity at surrounding points—or, according to our interpretation, how the distortion of space at any point fits on to the distortion at surrounding points. It is evidently an attempt to express the general laws of the strains in space and time which occur in Nature. If we are correct in our assumption that gravitation involves ***nothing more*** than strain of space-time,[2] so that its law expresses merely the relation between adjacent strains which holds by some natural necessity, clearly the strains which give the centrifugal force must obey the same general law. Here a very interesting point arises. We cannot reconcile the Newtonian law of gravitation with this condition. Newton's law and the law of centrifugal force are contradictory.

To put the matter another way, if we determine the strains by Newton's law, we get results closely agreeing with observation, provided Minkowski's space-time is used; but if we avail ourselves of our right to use a transformed space-time, the results no longer agree with observation. That means that Newton's law involves something which is not fully represented by strains, and so does not agree with our assumption. We must abandon either our assumption, or the famous law which has been accepted for more than

[1] Discourse delivered at the Royal Institution on Friday, February 1, by Prof. A. S. Eddington, F.R.S. Continued from p. 17.

[2] The idea is that matter represents a seam or nucleus of strain, and the strains at other points link themselves on according to laws inherent in the *continuum* and quite independent of the matter. The matter starts the strain, but does not *control* it as it goes outwards.

200 years, and find a new law of gravitation which will fall in with our requirements.

This amended law has been found by Einstein. It appears to be the only possible law that meets our requirements, and in the limited applications which come under practical observation is sufficiently close to the old law that has served so well. In practical applications the two laws are indistinguishable, except for one or two crucial phenomena to which reference will be made later. But in gravitational fields far stronger than any of which we have experience, and for bodies moving with velocities much greater than those of the planets, the difference would be considerable.

This idea of the distortion of space as the *modus operandi* of gravitation has led to a practical result—a new law of gravitation. It is not brought in as a hypothetical explanation of gravitation; if Einstein's theory is true, it is simply of the nature of an experimental fact.

If we draw a circle on a sheet of paper and measure the ratio of the circumference to the diameter, the result gives, if the experiment is performed accurately enough, the well-known number π, which has been calculated to 707 places of decimals. Now place a heavy particle at or near the centre and repeat the experiment; the ratio will be not exactly equal to π, but a little less. The experiment has not been performed, and is not likely to be performed, because the difference to be looked for is so small; but, if Einstein's theory is correct, that must be the result. The space around the heavy particle does not obey ordinary geometry; it is non-Euclidean. The change in its properties is not metaphysical, but something which, with sufficient care, could be measured. You can keep to Euclidean space if you like, and say that the measuring-rod has contracted or expanded according as it is placed radially or transversely to the gravitational force. That is all very well if the effect is small, but in a very intense gravitational field it would lead to ridiculous results like those we noticed in connection with the Michelson-Morley experiment—everything expanding or contracting as it changed position, and no one aware of any change going on. I think we have learnt our lesson that it is better to be content with the space of experience, whether it turns out to be Euclidean or not, and to leave to the mathematician the transformation of the phenomena into a space with more ideal properties.

This consequence of the new law of gravitation, though theoretically observable, is not likely to be put to any practical test either now or in the immediate future. But there are other consequences which just come within the range of refined observation, and so give an immediate practical importance to the new theory, which has indeed scored one very striking success. If we could isolate the sun and a single planet, then under the Newtonian law of gravitation the planet would revolve in an ellipse, repeating the same orbit indefinitely. Under the new law this is not quite true; the orbit is nearly an ellipse, but it does not exactly close up, and in the next revolution the planet describes a new ellipse in a slightly advanced position. In other words, the elliptic orbit slowly turns round in the same direction in which the planet is moving, so that after the lapse of many centuries the orbit will point in a different direction. The rate at which the orbit turns depends on the speed of motion of the planet in its orbit, so we naturally turn to the fastest moving planets, Mercury, Venus, and the earth, to see if the effect can be detected. Mercury moves at thirty miles a second, Venus at twenty-two, the earth at eighteen and a half. But there is a difficulty about Venus and the earth. Their orbits are nearly circular, and you cannot tell in which direction a circle is pointing. Mercury combines the favourable conditions of a high speed and a satisfactorily elongated orbit the direction of which at any time can be measured with considerable precision. It is found by observation that the orbit of Mercury is advancing at the rate of 574 seconds of arc a century. This is in great measure due to the attraction of the other planets, which are pulling the orbit out of shape and changing its position. The amount of this influence can be calculated very accurately, and amounts to 532 seconds per century. There is thus a difference of forty-two seconds a century unaccounted for; and this has for long been known as one of the most celebrated discordances between observation and gravitational theory in astronomy. It is thirty times greater than the probable error which we should expect from uncertainties in the observations and theory. There are other puzzling discordances, especially in connection with the motion of the moon; but the conditions in that case are more complicated, and I scarcely think they offer so direct a challenge to gravitational theory. Now Einstein's theory predicts that there will be a rotation of the orbit of Mercury additional to that produced by the action of the planets; and it predicts the exact amount—namely, that in one revolution of the planet the orbit will advance by a fraction of a revolution equal to three times the square of the ratio of the velocity of the planet to the velocity of light. We can work that out, and we find that the advance should be forty-three seconds a century—just about the amount required. Thus, whilst the Newtonian law leaves a discordance of more than forty seconds, Einstein's law agrees with observation to within a second or so.

Of course this superiority would be discounted if we could find some other application where the old Newtonian law had proved the better. But that has not happened. In all other cases the two laws agree so nearly that it has not been possible to discriminate between them by observation. The new law corrects the old where the old failed, and refrains from spoiling any agreement that already exists. The next best chance of applying the new theory is in the advance of the orbit of Mars; here Einstein's new law "gilds refined gold" by slightly improving an agreement which was already sufficiently good—a "wasteful and ridiculous excess," which is at any rate not unfavourable to the new theory.

There is another possibility of testing Einstein's theory, which it is hoped to carry out at the first opportunity. This relates to the action of gravitation on a ray of light. It is now known that electromagnetic energy possesses the property of inertia or mass, and probably the whole of the mass of ordinary matter is due to the electromagnetic energy which it contains. Light is a form of electromagnetic energy, and therefore must have mass—a conclusion which has been found true experimentally, because light falling on any object exerts a pressure just as a jet of water would. We ordinarily measure mass in pounds, and it is quite proper to speak of "a pound of light," just as we speak of a pound of tobacco. In case anyone should be thinking of going to an electric light company to buy a pound of light, I had better warn you that it is a rather expensive commodity. They usually prefer to sell it by a mysterious measure of their own, called the Board of Trade unit, and charge at least 3*d*. a unit. At that rate I calculate that they would let you have a pound of light for 141,615,000*l*. Fortunately, we get most of our light free of charge, and the sun showers down on the earth 160 tons daily. It is just as well we are not asked to pay for it.

But although light has mass, it does not follow that light has weight. Ordinarily, mass and weight are associated in a constant proportion, but whether this is so in the case of light can be settled only by experiment—by weighing light. It seems that it should be just possible to do this. If a beam of light passes an object which exerts a gravitational attraction, then, if it really has weight, it must drop a little towards the object. Its path will be bent just as the trajectory of a rifle bullet is curved owing to the weight of the bullet. The velocity of light is so great that there is only one body in the solar system powerful enough to make an appreciable bend in its path, namely, the sun. If we could see a star close up to the edge of the sun, a ray of light coming from the star would bend under its own weight, and the star would be seen slightly displaced from its true position. During a total eclipse stars have occasionally been photographed fairly close to the sun, and with care it should be possible to observe this effect. There is a magnificent opportunity next year when a total eclipse of the sun takes place right in the midst of a field of bright stars. This is the best opportunity for some generations, and it is hoped to send out expeditions to the line of totality to weigh light according to this method.

In any case, great interest must attach to an attempt to settle whether or not light has weight. But there is an additional importance, because it can be made a means of confirming or disproving Einstein's theory. On Einstein's theory light must certainly have weight, because mass and weight are viewed by it as two aspects of the same thing; but his theory predicts a deflection twice as great as we should otherwise expect. Apart from surprises, there seem to be three possible results:—(1) A deflection amounting to 1·75″ at the limb of the sun, which would confirm Einstein's theory; (2) a deflection of 0·83″ at the limb of the sun, which would overthrow Einstein's theory, but establish that light was subject to gravity; (3) no deflection, which would show that light, though possessing mass, has no weight, and hence that Newton's law of proportionality between mass and gravitation has broken down in another unexpected direction.

The purpose of Einstein's new theory has often been misunderstood, and it has been criticised as an attempt to explain gravitation. The theory does *not* offer any explanation of gravitation; that lies quite outside its scope, and it does not even hint at a possible mechanism. It is true that we have introduced a definite hypothesis as to the relation between gravitation and a distortion of space; but if that explains anything, it explains not gravitation, but space, *i.e.* the scaffolding constructed from our measures. Perhaps the position reached may be made clearer by another analogy. Let us picture the particle which describes a world-line as hurdle-racer in a field thickly strewn with hurdles. The particle in passing from point to point always takes the path of least effort, crossing the fewest possible hurdles; if the hurdles are uniformly distributed, corresponding with undistorted Minkowskian space, this will, of course, be a straight line. If the field is now distorted by a mathematical transformation such as an earthquake so that the hurdles become packed in some parts and spread out in others, the path of least effort will no longer be a straight line; but it is not difficult to see that it passes over precisely the same hurdles as before, only in their new positions. The gravitational field due to a particle corresponds with a more fundamental rearrangement of the hurdles, as though someone had taken them up and replanted them according to a law which expresses the law of gravitation. Any other particle passing through this part of the field follows the guiding rule of least effort, and curves its path, if necessary, so as to jump the fewest hurdles. Now, we have usually been under the impression that when we measured distances by physical experiments we were surveying the *field*, and the results could be plotted on a map; but it is now realised that we cannot do that. The field itself has nothing to do with our measurements; all we do is to count hurdles. If the only cause of irregularity of the hurdles were earthquakes (mathematical transformations), that would not make much difference, because we could still plot our counts of hurdles consistently as distances on a map; and the map would represent the original condition of the field with the hurdles uniformly spaced. But the more far-reaching rearrangement of hurdles by the gravitational field forces us to recognise that we are dealing with counts of hurdles and not with distances; because if we plot our measures on a map they will not close up. The number of hurdles in the circumference of a circle[3] will not be π times the number in the diameter; and when we try to draw on a map a circle the circumference of which is less than π times its diameter, we get into difficulties—at least in Euclidean space. This analogy brings out the point that the theory is an explanation of the real nature of our measures rather than of gravitation. We offer no explanation why the particle always takes the path of least effort—perhaps, if we may judge by our own feelings, that is so natural as to require no explanation. More seriously, we know that in consequence of the undulatory theory of light, a ray traversing a heterogeneous medium always takes the path of least time; and one can scarcely resist a vague impression that the course of a material particle may be the ray of an undulation in five dimensions. What concerns gravitation more especially is that we have offered no explanation of the linkages by which the hurdles rearrange themselves on a definite plan when disturbed by the presence of a gravitating particle; that is a point on which a mechanical theory of gravitation ought to throw light.

From the constant of gravitation, together with the other fundamental constants of Nature—the velocity of light and the quantum of action—it is possible to form a new fundamental unit of length. This unit is 7×10^{-28} cm. It seems to be inevitable that this length must play some fundamental part in any complete interpretation of gravitation. (For example, in Osborne Reynolds's theory of matter this length appears as the mean free-path of the granules of his medium.) In recent years great progress has been made in knowledge of the excessively minute; but until we can appreciate details of structure down to the quadrillionth or quintillionth of a centimetre, the most sublime of all the forces of Nature remains outside the purview of the theories of physics.

UNIVERSITY AND EDUCATIONAL INTELLIGENCE.

CAMBRIDGE.—The syndicate appointed to consider the Previous examination has issued its report. The recommendations involve changes which, if passed by the Senate, will greatly alter the present character of the examination. The syndicate advises the discontinuance of Greek as a compulsory subject, and recommends that every candidate shall be required to take at least one paper in natural science. It proposes that the examination should be in three parts: (1) Languages; (2) mathematics and natural science; (3) English subjects. In part (1) it is recommended that Latin should continue to be a compulsory subject, and that a candidate should be required to take one other foreign language, namely, Greek, French, German, or Spanish.

[3] A circle would naturally be defined as a curve such that the number of hurdles (counted along the path of least effort) between any point on it and a fixed point called the centre is constant. To make the vague analogy more definite, we may suppose that the hurdles are pivoted, and swing round automatically to face the jumper; he is not allowed to dodge them, *i.e.* to introduce into his path sinuosities comparable with the lengths of the hurdles

Chapter 4
George B. Jeffery (1921): The Field of an Electron on Einstein's Theory of Gravitation

George B. Jeffery (1921). The Field of an Electron on Einstein's
Theory of Gravitation. *Proceedings of the Royal Society of London*. Series A 99 (697),
123–134.

The Field of an Electron on Einstein's Theory of Gravitation.

By G. B. JEFFERY, M.A., D.Sc., Fellow of University College, London.

(Communicated by Prof. L. N. G. Filon, F.R.S. Received December 16, 1920,— Revised January 11, 1921.)

Einstein's Generalised Theory of Relativity has accomplished notable results in the region of astronomical mechanics, but for the moment it seems difficult to go further in this direction until some further progress is made with the theory of the solution of the field equations. Practically all the consequences of the theory which have been established so far are obtained from the solution of the equations corresponding to a single isolated singularity. The *exact* solutions corresponding to two isolated singularities are urgently required for the further exploration of the theory. The approximate solutions which have been put forward by De Sitter,* Dröste† and Einstein,‡ though valuable in default of exact solutions, are apt to be misleading and perhaps to exclude effects of far-reaching theoretical importance such as radiation. Meanwhile it is well to examine the consequences of the theory at the other extreme of the realm of physical science, in its relation to atomic phenomena. Here it seems no longer justifiable to consider gravitation and electricity separately. In these microscopic phenomena, where we are free from the effects of averaging, mass and charge seem to be inextricably connected. They exist in certain definite combinations, possibly in only two combinations as electrons and hydrogen nuclei. The gravitational field corresponding to such a charged particle has been investigated by Nordström§ by the application of the calculus of variations to the Hamiltonian function of the combined fields. Nordström's work was not brought to our notice until after the present paper was written and we had obtained the same result by direct solution of the field equations. As the methods employed may perhaps be more familiar to English students of the subject it has been thought well to give a brief outline of this proof as an alternative to Nordström's in the first section of this paper.

* 'Monthly Notices of the Royal Astronomical Society,' vol. 72, p. 155 (1916).

† 'Proc. Acad. Amsterdam,' vol. 19, p. 447 (1916).

‡ 'Berlin, Sitzungsberichte,' 1916, p. 688.

§ "On the Energy of the Gravitational Field in Einstein's Theory," 'Proc. Ac. Amsterdam,' vol. 20, p. 1236 (1918).

§ 1. *The Solution of the Field Equations.*

The gravitational field is defined by the symmetrical covariant tensor $g_{\mu\nu}$, where

$$ds^2 = \Sigma\Sigma\,(\mu, \nu)\, g_{\mu\nu}\, dx_\mu\, dx_\nu \qquad (\mu, \nu = 1, 2, 3, 4).$$

where the variables of summation are enclosed in brackets immediately following the signs of summation.

For a symmetrical static field we may, following Eddington* and Schwartzchild, take polar co-ordinates r, θ, ϕ, $\tau = ct$ and

$$ds^2 = -e^\lambda\, dr^2 - r^2\, d\theta^2 - r^2 \sin^2\theta\, d\phi^2 + e^\nu\, d\tau^2, \tag{1}$$

where λ, ν are functions of r only, so that $g_{\mu\nu} = 0$ if $\mu \neq \nu$ and

$$g_{11} = -e^\lambda, \quad g_{22} = -r^2, \quad g_{33} = -r^2 \sin^2\theta, \quad g_{44} = e^\nu, \tag{2}$$

The determinant of $g_{\mu\nu}$ is given by

$$g = -r^4 \sin^2\theta\, e^{\lambda+\nu}. \tag{3}$$

and the associated contravariant tensor by

$$g^{\mu\mu} = 1/g_{\mu\mu} \qquad \text{and} \qquad g^{\mu\nu} = 0 \text{ for } \mu \neq \nu. \tag{4}$$

The Christoffel symbols, defined by

$$\{\lambda\mu, \nu\} = \tfrac{1}{2}\Sigma\,(\alpha)\, g^{\nu\alpha}\left(\frac{\partial g_{\lambda\alpha}}{\partial x_\mu} + \frac{\partial g_{\mu\alpha}}{\partial x_\lambda} - \frac{\partial g_{\lambda\mu}}{dx_\alpha}\right),$$

are then as given by Eddington,† accents denoting differentiation with respect to r,

$$\left.\begin{aligned}
&\{11, 1\} = \tfrac{1}{2}\lambda', \qquad \{12, 2\} = \{13, 3\} = 1/r, \qquad \{14, 4\} = \tfrac{1}{2}\nu',\\
&\{22, 1\} = -re^{-\lambda}, \qquad \{23, 3\} = \cot\theta,\\
&\{33, 1\} = -r\sin^2\theta e^{-\lambda}, \qquad \{33, 2\} = -\sin\theta\cos\theta,\\
&\{44, 1\} = \tfrac{1}{2}\nu' e^{\nu-\lambda},
\end{aligned}\right\}, \tag{5}$$

the rest vanishing identically.

The components of Einstein's contracted tensor $G_{\mu\nu}$ vanish with the exception of

$$\left.\begin{aligned}
G_{11} &= \tfrac{1}{2}\nu'' - \tfrac{1}{4}\lambda'\nu' + \tfrac{1}{4}\nu'^2 - \lambda'/r.\\
G_{22} &= e^{-\lambda}\{1 + \tfrac{1}{2}r(\nu' - \lambda')\} - 1.\\
G_{33} &= \sin^2\theta\,[e^{-\lambda}\{1 + \tfrac{1}{2}r(\nu' - \lambda')\} - 1]\\
G_{44} &= -e^{\nu-\lambda}\{\tfrac{1}{2}\nu'' - \tfrac{1}{4}\lambda'\nu' + \tfrac{1}{4}\nu'^2 + \nu'/r\}.
\end{aligned}\right\}. \tag{6}$$

* 'Report on Relativity to the Physical Society,' second edition, p. 44. Eddington's Methods have been closely followed in this paper, and references to this source will be quoted as 'Report.'

† 'Report,' p. 45.

The electromagnetic field is defined* by the covariant vector

$$\kappa_\mu \equiv (-\mathrm{F},\ -\mathrm{G},\ -\mathrm{H},\ \Phi)$$

where F, G, H, are the components of the vector potential and Φ is the scalar potential. For static symmetry about the origin $\kappa_2 = \kappa_3 = 0$, and κ_1, κ_4 are functions of r only.

Denoting the covariant derivative of κ_μ by $\kappa_{\mu\sigma}$ the electromagnetic covariant tensor is defined by

$$\mathrm{F}_{\mu\sigma} = \kappa_{\mu\sigma} - \kappa_{\sigma\mu},$$

and in our case all the components vanish with the exception of

$$\mathrm{F}_{14} = -\kappa_4', \qquad \mathrm{F}_{41} = \kappa_4', \tag{7}$$

the accents as before denoting differentiation with respect to r.

The associated mixed and contravariant tensors are defined respectively by

$$\mathrm{F}_\mu{}^\sigma = \Sigma\,(\alpha)\, g^{\sigma\alpha}\,\mathrm{F}_{\mu\alpha}, \qquad \mathrm{F}^{\mu\sigma} = \Sigma\Sigma\,(\alpha,\beta)\, g^{\mu\alpha} g^{\sigma\beta}\,\mathrm{F}_{\alpha\beta}, \tag{8}$$

and in our case the only surviving components are

$$\mathrm{F}_4{}^1 = g^{11}\mathrm{F}_{41} = -e^{-\lambda}\kappa_4', \qquad \mathrm{F}_1{}^4 = g^{44}\mathrm{F}_{14} = -e^{-\nu}\kappa_4' \tag{9}$$

and

$$\mathrm{F}^{14} = g^{11}g^{44}\,\mathrm{F}_{14} = e^{-(\lambda+\nu)}\,\kappa_4' = -\mathrm{F}^{41}. \tag{10}$$

In the absence of charge, Maxwell's equations of the electromagnetic field are expressed† by the vanishing of the contracted covariant derivative of $\mathrm{F}^{\mu\sigma}$,

$$\mathrm{F}_\sigma{}^{\mu\sigma} \equiv \frac{1}{\sqrt{(-g)}}\,\Sigma\,(\sigma)\,\frac{\partial}{\partial x_\sigma}\left[\sqrt{(-g)}\,\mathrm{F}^{\mu\sigma}\right]$$

in Lorentz units.

These lead to a single equation in κ_4,

$$\frac{\partial}{\partial r}\left(r^2\kappa_4' e^{-\frac{1}{2}(\lambda-\nu)}\right) = 0. \tag{11}$$

The mixed electromagnetic energy tensor is defined by

$$\mathrm{E}_\sigma{}^\mu = -\Sigma\,(\alpha)\,\mathrm{F}_{\sigma\alpha}\,\mathrm{F}^{\mu\alpha} + \tfrac{1}{4} g_\sigma{}^\mu\,\Sigma\Sigma\,(\alpha,\beta)\,\mathrm{F}^{\alpha\beta}\,\mathrm{F}_{\alpha\beta}.$$

In our case the only surviving components are

$$\left.\begin{aligned} \mathrm{E}_1{}^1 = \mathrm{E}_4{}^4 &= \tfrac{1}{2}\kappa_4'^2\, e^{-(\lambda+\nu)} \\ \mathrm{E}_2{}^2 = \mathrm{E}_3{}^3 &= -\tfrac{1}{2}\kappa_4'^2\, e^{-(\lambda+\nu)} \end{aligned}\right\}. \tag{12}$$

The associated covariant energy tensor

$$\mathrm{E}_{\mu\sigma} = \Sigma\,(\alpha)\, g_{\alpha\sigma}\,\mathrm{E}_\mu{}^\alpha$$

vanishes except for the components

$$\left.\begin{aligned} \mathrm{E}_{11} &= -\tfrac{1}{2}\kappa_4'^2 e^{-\nu}, & \mathrm{E}_{22} &= \tfrac{1}{2} r^2 \kappa_4'^2 e^{-(\lambda-\nu)} \\ \mathrm{E}_{33} &= \tfrac{1}{2} r^2 \sin^2\theta\, \kappa_4'^2 e^{-(\lambda+\nu)}, & \mathrm{E}_{44} &= \tfrac{1}{2}\kappa_4'^2 e^{-\lambda} \end{aligned}\right\}, \tag{13}$$

* 'Report,' p. 76.

† 'Report,' p. 77.

and the associated scalar

$$\mathrm{E} = \Sigma\Sigma\,(\mu\sigma)\, g^{\mu\sigma}\, \mathrm{E}_{\mu\sigma} = \Sigma\,(\alpha,\,\mu)\, g_{\alpha}{}^{\mu}\, \mathrm{E}_{\mu}{}^{\alpha} = \mathrm{E}_1{}^1+\mathrm{E}_2{}^2+\mathrm{E}_3{}^3+\mathrm{E}_4{}^4 = 0. \quad (14)$$

Omitting the terms representing the density and motion of matter, Einstein's equations of the gravitational field are

$$\mathrm{G}_{\mu\sigma} = -8\pi\kappa c^{-4}\,(\mathrm{E}_{\mu\sigma} - \tfrac{1}{2} g_{\mu\sigma}\,\mathrm{E}),$$

where κ is the constant of gravitation and c the velocity of light

These give

$$\mathrm{G}_{11} = 4\pi\kappa c^{-4}\kappa_4'^2 e^{-\nu},$$

$$\mathrm{G}_{22} = -4\pi\kappa c^{-4} r^2 \kappa_4'^2 e^{-(\lambda+\nu)},$$

$$\mathrm{G}_{33} = -4\pi\kappa c^{-4} r^2 \sin^2\theta\, \kappa_4'^2\, e^{-(\lambda+\nu)},$$

$$\mathrm{G}_{44} = -4\pi\kappa c^{-4}\kappa_4'^2 e^{-\lambda},$$

the remaining components of $\mathrm{G}^{\mu\sigma}$ vanishing.

Hence from (6) and (15) we have

$$\tfrac{1}{2}\nu'' - \tfrac{1}{4}\lambda'\nu' + \tfrac{1}{4}\nu'^2 - \lambda' r = 4\pi\kappa c^{-4}\kappa_4'^2 e^{-\nu}, \quad (16)$$

$$e^{-\lambda}\{1+\tfrac{1}{2} r(\nu'-\lambda')\}-1 = -4\pi\kappa c^{-4} r^2 \kappa_4'^2 e^{-(\lambda+\nu)}, \quad (17)$$

$$\sin^2\theta\,[e^{-\lambda}\{1+\tfrac{1}{2} r(\nu'-\lambda')\}-1] = -4\pi\kappa c^{-4} r^2 \sin^2\theta\,\kappa_4'^2 e^{-(\lambda+\nu)}, \quad (18)$$

$$\tfrac{1}{2}\nu'^1 - \tfrac{1}{4}\lambda'\nu' + \tfrac{1}{4}\nu'^2 + \nu'/r = 4\pi\kappa c^{-4}\kappa_4'^2 e^{-\nu}. \quad (19)$$

Equations (17) and (18) are clearly identical, while (16) and (19) give $\lambda'+\nu' = 0$ or $\lambda+\nu = const.$ If the gravitational field disappears at a great distance from the origin we may take $\lambda+\nu = 0$.

Equation (11) then gives

$$\kappa_4 = \frac{\epsilon}{4\pi r}, \quad (20)$$

where ϵ is a constant which will presently be identified with the electric charge in Lorentz units, and a second constant of integration has been put equal to zero without loss of generality.

Equation (17) then gives

$$e^{\nu}\,(1+r\nu') = 1 - \frac{\kappa\epsilon^2}{4\pi c^4 r^2},$$

or writing $e^{\nu} = \gamma$,

$$\gamma = 1 - \frac{2\kappa m}{c^2 r} + \frac{\kappa\epsilon^2}{4\pi c^4 r^2}, \quad (21)$$

where m is a constant of integration which will subsequently be seen to be the mass. It may now be verified that equations (16) and (19) are satisfied identically.

We have therefore as the line element of the space surrounding a point which is a singularity of *both* the electric and gravitational fields

$$ds^2 = -\gamma^{-1} dr^2 - r^2 d\theta^2 - r^2\sin^2\theta\, d\phi^2 + \gamma c^2 dt^2, \quad (22)$$

which is the result given by Nordström.

§2. *The Motion of a Particle in the Field of a Charged Nucleus.*

We will investigate the motion of a particle having both charge and inertia mass in the field which we have examined in the previous section. The mass and charge of the particle will be supposed so small that, while the motion of the particle is determined by the field, the field itself is not appreciably affected by the motion of the particle. This is a serious limitation, but it is the best that we can do in default of the solution corresponding to two singularities. It neglects all effects due to radiation and to the reaction upon the nucleus. If the nucleus is itself an electron, it is difficult to imagine any physical phenomenon to which the analysis can apply, unless indeed we regard the electron as composed of still smaller elements in a disposition which is stable for the particular values of the electronic charge, mass and radius. The analysis will, however, give a good approximation to the motion of an electron in the neighbourhood of an atomic nucleus, for then the mass of the nucleus is at least about a thousand times that of the electron, and, while the charges may be of the same order, it will appear that the terms in $g_{\mu\nu}$ which depend on the charge are in all practical cases small compared with those depending upon the mass.

The "force" acting on a charged particle is given by* the covariant vector

$$k_\sigma = \mathrm{F}_{\sigma\mu}\,\mathrm{J}^\mu,$$

where J^μ is the charge-current vector which, for a single particle of charge ϵ' is

$$\epsilon'\left(\frac{dx_1}{ds},\ \frac{dx_2}{ds},\ \frac{dx_3}{ds},\ \frac{dx_4}{ds}\right).$$

The associated contravariant vector is

$$k^\sigma = \Sigma\,(\alpha)\,g^{\alpha\sigma}\,k_\alpha. \tag{23}$$

The contravariant acceleration vector† is

$$\mathrm{A}^\sigma = \frac{d^2x_\sigma}{ds^2} + \Sigma\Sigma\,(\alpha,\beta)\,\{\alpha\beta,\sigma\}\,\frac{dx_\alpha}{ds}\,\frac{dx_\beta}{ds}. \tag{24}$$

We will take as the equations of motion of our particle

$$m'\,\mathrm{A}^\sigma = -\frac{1}{c^2}\,k^\sigma, \tag{25}$$

where m' is the mass of the particle.

In rectangular co-ordinates $x, y, z, \tau = ct$ in a non-gravitational field, the Christoffel symbols vanish, and if v is the velocity of the particle

$$\frac{d}{ds} = \frac{1}{\sqrt{(c^2-v^2)}}\,\frac{d}{dt}.$$

* 'Report,' p. 78.

† 'Report,' p. 48.

Remembering that the electric and magnetic vectors (E, B) are given by

$$\mathrm{B} = curl\,(\mathrm{F, G, H}), \qquad \mathrm{E} = -grad\Phi - \frac{1}{c}\frac{\partial}{\partial t}(\mathrm{F, G, H}),$$

the equations (25) are easily reduced to

$$\frac{d}{dt}\left\{\frac{m'}{\sqrt{(1-v^2/c^2)}}\frac{dx}{dt}\right\} = \epsilon'\left(\mathrm{E}_x + \frac{1}{c}(v_y\mathrm{B}_z - v_z\mathrm{B}_y)\right)$$

and two similar equations, with

$$\frac{d}{dt}\left\{m'c^2\left(\frac{1}{\sqrt{(1-v^2/c^2)}}-1\right)\right\} = \epsilon'(\mathrm{E}_x v_x + \mathrm{E}_y v_y + \mathrm{E}_z v_z).$$

These are the equations of motion on the restricted relativity theory, and reduce to the classical equations for small velocities. Equations (25), therefore, hold in the absence of gravitation, and being of the necessary tensor form they will, by the equivalence hypothesis, be true for any system of co-ordinates and in a gravitational field.

In our co-ordinates

$$k_1 = \frac{\epsilon\epsilon' c}{4\pi r^2}\frac{dt}{ds}, \quad k_2 = k_3 = 0, \quad k_4 = -\frac{\epsilon\epsilon'}{4\pi r^2}\frac{dr}{ds}$$

and hence

$$k^1 = -\frac{\epsilon\epsilon' c e^{\nu}}{4\pi r^2}\frac{dt}{ds}, \quad k^2 = k^3 = 0, \quad k^4 = -\frac{\epsilon\epsilon' e^{-\nu}}{4\pi r^2}\frac{dr}{ds}.$$

Substituting from (5) and (24) in (25) we have

$$\frac{d^2r}{ds^2} - \tfrac{1}{2}\nu'\left(\frac{dr}{ds}\right)^2 - re^{\nu}\left(\frac{d\theta}{ds}\right)^2 - r\sin^2\theta\, e^{\nu}\left(\frac{d\phi}{ds}\right)^2 + \tfrac{1}{2}c^2\nu' e^{2\nu}\left(\frac{dt}{ds}\right)^2 = \frac{\epsilon\epsilon' e^{\nu}}{4\pi c m' r^2}\frac{dt}{ds}, \quad (26)$$

$$\frac{d^2\theta}{ds^2} + \frac{2}{r}\frac{dr}{ds}\frac{d\theta}{ds} - \sin\theta\cos\theta\left(\frac{d\phi}{ds}\right)^2 = 0, \quad (27)$$

$$\frac{d^2\phi}{ds^2} + \frac{2}{r}\frac{dr}{ds}\frac{d\phi}{ds} + 2\cot\theta\,\frac{d\theta}{ds}\frac{d\phi}{ds} = 0, \quad (28)$$

$$\frac{d^2t}{ds^2} + \nu'\frac{dr}{ds}\frac{dt}{ds} = \frac{\epsilon\epsilon' e^{-\nu}}{4\pi c^3 m' r^2}\frac{dr}{ds}. \quad (29)$$

For motion in the equatorial plane we may put $\theta = \frac{1}{2}\pi$, and (27) is then identically satisfied. From (28) and (29) we then have

$$r^2\frac{d\phi}{ds} = h, \qquad \gamma\frac{dt}{ds} = n - \frac{\epsilon\epsilon'}{4\pi c^3 m' r}, \quad (30)$$

where h and n are constants. Eliminating ϵ, ϵ' between (26) and (29) we obtain a third integral

$$\frac{1}{\gamma}\left(\frac{dr}{ds}\right)^2 + r^2\left(\frac{d\phi}{ds}\right)^2 - c^2\gamma\left(\frac{dt}{ds}\right)^2 + 1 = 0, \quad (31)$$

where γ is given by (21).

Considering a particle at rest in three dimensional space so that $\frac{dr}{ds} = \frac{d\phi}{ds} = 0$, we obtain from (26) and (31)

$$m' \frac{d^2r}{dt^2} = -\kappa \frac{mm'}{r^2} + \frac{\epsilon\epsilon'}{4\pi r^2}$$

where terms of higher order than r^{-2} have been neglected. Hence this is the equation of motion of a particle at great distances from the origin, and our interpretation of the constants m, ϵ as the mass and charge of the nucleus respectively, is confirmed.

We have used Lorentz units in order to maintain symmetry among the components of the various tensors employed. We may now revert to ordinary electrostatic units by omitting the factor 4π in (21) and (30). We then have as our equations of motion,

$$\frac{1}{\gamma}\left(\frac{dr}{ds}\right)^2 + r^2\left(\frac{d\phi}{ds}\right)^2 - c^2\gamma\left(\frac{dt}{ds}\right)^2 + 1 = 0, \tag{32}$$

with

$$r^2 \frac{d\phi}{ds} = h, \tag{33}$$

and

$$\gamma \frac{dt}{ds} = n - \frac{\epsilon\epsilon'}{c^2 m' r}, \tag{34}$$

where

$$\gamma = 1 - \frac{2\kappa m}{c^2 r} + \frac{\kappa\,\epsilon^2}{c^4 r^2}. \tag{35}$$

If we eliminate t, s from (32) by means of (33) and (34), and write $u = 1/r$, we obtain as the geometrical equation of the orbits

$$h^2\left(\frac{du}{d\phi}\right)^2 = c^2\left(n - \frac{\epsilon\epsilon'}{c^3 m'}u\right)^2 - (1 + h^2u^2)\left(1 - \frac{2\kappa m}{c^2}u + \frac{\kappa\epsilon^2}{c^4}u^2\right). \tag{36}$$

We will reserve a detailed investigation of these orbits for a future communication. We may, however, point out that it follows from (36) that u, and therefore r, can be expressed as an elliptic function of ϕ. Since the coefficients of (36) are real and the right-hand side is necessarily positive, this elliptic function has a real period. Hence the orbit is periodic in the sense that r is a periodic function of ϕ, although the period will not in general be 2π. The orbit will, therefore, not be closed, and there will be a rotation of the apse lines. This periodicity of the orbit is due to the fact that we have taken no account of the modification in the field due to the moving particle. A more complete solution, which included the effects of both particles, would probably lead to orbits which would not generally be periodic. This would correspond to the fact pointed out by Einstein, that moving masses imply radiation of gravitational energy.

Einstein seems to infer from this that his law of gravitation must be subject to the restrictions of the Quantum theory. He says* :—

"Gleichwohl müssten die Atome zufolge der inneratomischen Elektronenbewegung nicht nur electromagnetische, sondern auch Gravitationsenergie ausstrahlen, wenn auch in winzigem Betrage. Da dies in Wahrheit in der Natur nicht zutreffen dürfte, so scheint es, dass die Quantentheorie nicht nur die Maxwellsche Elektrodynamik, sondern auch die neue Gravitationstheorie wird modifizieren müssen."

Einstein may be correct in his speculation, but is there not another possibility? An investigation of the motion of a particle of infinitesimal mass in the field of a particle of finite mass has led us to equation (36) and orbits which are always strictly periodic. The complete solution of the problem of two bodies might well lead to an equation corresponding to (36) and to orbits which are not in general periodic, but which *may* in certain circumstances be periodic. These would give the "quantised" orbits. If it should then appear that the non-periodic orbits tended to asymptote down to the periodic orbits, and we could trace the radiation emitted in the process, the whole secret of the Quantum hypothesis would be laid bare. This is, of course, mere speculation, but it may serve to show the extreme importance of obtaining the exact solution of Einstein's equations corresponding to *two* point singularities.

§ 3. *The Conception of a Point Electron.*

We observe from (22) that the effect of the singularity upon the gravitational field is given by the deviation of γ from its value at infinity, namely unity. As we approach the singularity from infinity γ steadily decreases to a minimum at $r = \epsilon^2/mc^2$ and then increases and finally reaches a positive infinity at the singularity. It is equal to its value at infinity when

$$r = \epsilon^2/2mc^2. \tag{37}$$

For an electron we may take $\epsilon = 3 \times 10^{-10}$ C.G.S. electrostatic units, $m = 10^{-27}$ grm. and $c = 3 \times 10^{10}$ cm./sec. and (37) then gives $r = 0{\cdot}5 \times 10^{-13}$ which is of the same order as the usually accepted value of the radius of an electron. It would be easy to attach too much importance to this numerical result. Firstly, it is not clear that the equality of γ, and, therefore, of the $g_{\mu\nu}$, at two different places has any precise physical meaning. It probably depends upon the mode of measuring r, which in this theory is largely at our disposal. Secondly, the third term in (35) represents the effect of electromagnetic energy upon the gravitational field and by equating γ to unity we

* 'Berlin Sitzungsberichte,' 1916, p. 696.

have obtained a relation between mass and electromagnetic energy. Such a relation is given by the ordinary electromagnetic theory of mass, and in fact as far as we are aware the only determination of the radius of an electron is by means of a formula for the radius in terms of the mass and charge, which differs from (37) only by a numerical factor which depends upon the distribution of charge in the electron. It is, therefore, not surprising that we should obtain a radius of the same order of magnitude. At the same time the present point of view suggests a new meaning for the "radius" of an electron. On the older theory of electromagnetic mass we were forced to regard an electron as a body of small but finite dimensions; it could not be a mere *point* singularity of the field, as the mass of such a point charge would be infinite.

In the same way we cannot have a point singularity on Einstein's theory if we neglect the effects of electric charge, for in this case $\gamma = 1 - 2\kappa m/c^2 r$ and $\gamma = 0$ for $r = 2\kappa m/c^2$. Hence at $r = 0$, $g_{11} = 0$, $g_{44} = \infty$ and at $r = 2\kappa m/c^2$ $g_{11} = \infty$, $g_{44} = 0$. We have, therefore, not a solution with a single point singularity, but a solution with a point singularity surrounded by a spherical surface of singularity.

On the other hand if we include the effects of charge, we see from (35) that γ has no zeros or infinities other than $r = 0$, if $\epsilon^2/m^2 > \kappa$, *i.e*, $> 0{\cdot}67 \times 10^{-7}$ C.G.S. units. This condition is amply satisfied, both in the case of the electron and the hydrogen nucleus. We may therefore, if we wish, regard these as true point singularities of the field. As we approach such a point very closely γ becomes very large and the field is profoundly modified. As a measure of the dimensions of this region we may take the radius of a sphere on the surface of which the g's have the same value as at infinity. It is this radius which we have calculated above as the radius of the electron.

§ 4. *The Effect of an Electron upon Radiation in its Field.*

According to Einstein's theory, a ray of light passing near to an attracting mass is deflected, and it appears that the deflection would be very great for a ray passing very close to the attracting mass. Indeed, it would seem to be not impossible that a ray which passed sufficiently close to an attracting particle might be so strongly deflected that it would be permanently entrapped by the particle. If this should prove to be a legitimate deduction from Einstein's theory, it might have far-reaching consequences on our views as to the nature of the electron. It will be shown, however, that if we allow for both the charge and the mass of the electron, even though we regard it as a point singularity, no such result is possible. The deflection is always finite. As the distance of nearest approach diminishes, the deflection increases up to a maximum and then diminishes, becomes negative, and finally approaches $-\pi$, so that the ray is returned very nearly along its original direction.

For the path of a ray of light, we have $ds = 0$, which, for a ray in the equatorial plane $\theta = \frac{1}{2}\pi$, gives

$$\gamma c^2 dt^2 = \gamma^{-1} dr^2 + r^2 d\phi^2. \tag{38}$$

We have also to make the integral $\int dt$ stationary for given initial and final points. From (38), we have

$$\int c\,dt = \int \frac{1}{\gamma}\left\{1+\gamma r^2\left(\frac{d\phi}{dr}\right)^2\right\}^{\frac{1}{2}} dr.$$

Hence

$$\begin{aligned}\delta\int c\,dt &= \int r^2\frac{d\phi}{dr}\left\{1+\gamma r^2\left(\frac{d\phi}{dr}\right)^2\right\}^{-\frac{1}{2}}\frac{d}{dr}(\delta\phi)\,dr \\ &= \left| r^2\frac{d\phi}{dr}\left\{1+\gamma r^2\left(\frac{d\phi}{dr}\right)^2\right\}^{-\frac{1}{2}}\delta\phi \right| \\ &\qquad -\int\frac{d}{dr}\left[r^2\frac{d\phi}{dr}\left\{1+\gamma r^2\left(\frac{d\phi}{dr}\right)^2\right\}^{-\frac{1}{2}}\right]\delta\phi\,dr.\end{aligned}$$

The expression to be taken between limits vanishes, since $\delta\phi$ vanishes at the extreme points of the path, and the integral vanishes if

$$r^4\left(\frac{d\phi}{dr}\right)^2 = p^2\left\{1+\gamma r^2\left(\frac{d\phi}{dr}\right)^2\right\}$$

where p^2 is a constant.

Writing $u = 1/r$, and substituting the value of γ from (35), this may be written

$$\left(\frac{du}{d\phi}\right)^2 = \frac{1}{p^2} - u^2 + \frac{2\kappa m}{c^2}u^3 - \frac{\kappa\epsilon^2}{c^4}u^4, \tag{39}$$

from which it may be shown that p is the length of the perpendicular drawn from the origin on to the asymptotes of the path of the ray.

The right-hand side of (39) is a quartic in u, and hence u may be expressed as an elliptic function of ϕ. This elliptic function could be worked out explicitly if for any reason it should become desirable to investigate the exact form of the rays. The analysis is however heavy, and we can obtain all the information we require by other methods.

Denoting the right-hand side of (39) by $f(u)$, we have

$$f'(u) = -2u\left(1 - \frac{3\kappa m}{c^2}u + \frac{2\kappa\epsilon^2}{c^4}u^2\right).$$

Inserting the values of m and ϵ given in the last section, it will be seen that the quadratic on the right-hand side has no real roots in the case of the electron or the hydrogen nucleus, and, therefore, the only real root of

$f'(u) = 0$ is $u = 0$. Hence $f(u) = 0$ has at most two real roots. It then appears from (39) that, for real rays $p^2 > 0$, so that all real rays have real asymptotes, and it is not possible for light to circle in a closed "light orbit" round the electron. If $p^2 > 0$, it follows that $f(u) = 0$ has one real positive root and one real negative root. Let $u = \lambda$ be the positive root then, if the angle between the asymptotes is 2α, we have

$$\alpha = \int_0^\lambda \left\{\frac{1}{p^2} - u^2 + \frac{2\kappa m}{c^2} u^3 - \frac{\kappa\epsilon^2}{c^4} u^4\right\}^{-\frac{1}{2}} du.$$

Writing $u = \lambda v$, and remembering that $f(\lambda) = 0$, this gives

$$\alpha = \int_0^1 (1-v)^{-\frac{1}{2}} \left\{1 + v - \frac{2\kappa m}{c^2} \lambda (1+v+v^2) + \frac{\kappa\epsilon^2}{c^4} \lambda^2 (1+v)(1+v^2)\right\}^{-\frac{1}{2}} dv. \quad (40)$$

For rays which approach the electron very closely, p is very small and λ is very large, and, for sufficiently large values of λ (40), approximates to

$$\alpha \doteqdot \frac{1}{\lambda} \frac{c^2}{\kappa^{\frac{1}{2}}\epsilon} \int_0^1 \frac{dv}{\sqrt{(1-v^4)}},$$

which tends to zero as λ increases. Hence rays which, if undisturbed, would pass very close indeed to the electron, are reflected back practically along their original direction.

For larger values of p, we may conveniently refer back to equation (39). Differentiating with respect to ϕ, we obtain

$$\frac{d^2u}{d\phi^2} + u = \frac{3\kappa m}{c^2} u^2 - \frac{2\kappa\epsilon^2}{c^4} u^3. \quad (41)$$

Allowing for the difference of the units employed, this agrees with the usual equation for the deflection of a ray of light, except for the presence of the last term, which gives the influence of the electric field.

The equation (41) may be integrated by successive approximation in the usual way if u is so small that the terms on the right-hand side are small compared with u. In this way we obtain as the total deflection of the ray

$$\frac{4\kappa m}{c^2 p} - \frac{3\pi\kappa\epsilon^2}{4c^4p^2}. \quad (42)$$

The first term is the well-known Einstein deflection, and the second term shows that the deflection is reduced by the electric field. This approximation will hold for sufficiently large values of p and we have already examined the case of very small values of p. The intermediate cases may be studied by means of (39) if need should arise.

§ 5. *The Influence of the Sun's Electric Field on the Crucial Phenomena of Einstein's Theory.*

In accordance with these results the deflection of a ray of light by the Sun and the predicted displacement of the lines of the Solar spectrum will be influenced by the Sun's electric field.

The ratio of the wave-length of the radiation from a terrestrial atom to that of the radiation from a similar atom in the Sun is, in accordance with the usually accepted theory, given by $\sqrt{\gamma}$. Now it appears from (35) that the effect of any electric field which the Sun may have, whatever its sign, will tend to counteract the effect of the Sun's mass, and we may inquire what strength of field would be necessary to produce exact compensation. For this purpose we must have $\gamma = 1$ at the Sun's surface, or from (35),

$$\epsilon^2 = 2mc^2r.$$

Such a charge on the Sun, if distributed symmetrically, would give at the Sun's surface a field of intensity

$$c\sqrt{\frac{2m}{r^3}} \text{ c.g.s. units} \quad \text{or} \quad c^2\sqrt{\frac{2m}{r^2}}\, 10^{-8} \text{ volts/cm.}$$

The quantity under the square root being of the same order as the Sun's mean density, this requires a potential gradient of the order of 10^{13} volts. per cm., which is far too high to be admissible. The solution to the vexed question of the displacement of the spectrum lines is not to be found in this direction.

In the same way it may be shown that, for any possible value of the Sun's electric field, the second term in (42) is very small compared with the first, and the electric field will not exert any measurable effect on the bending of the rays of light in the gravitational field of the Sun.

Shifting Boundaries between Quantum Theory and Relativity

Chapter 5
Where to Start? First Interactions between Wave Mechanics and General Relativity

Alexander Blum

In the years 1925 to 1927, the old quantum theory was replaced by the newly developed theory of quantum mechanics, which grew out of the matrix mechanics of Heisenberg, Born, and Jordan and the wave mechanics of Schrödinger. Both of these theories were initially formulated entirely non-relativistically. But it was clear from the outset that contact would have to be made with the special theory of relativity for two important reasons: On the one hand, the mechanics would have to be complemented with a quantum electrodynamics (QED), in order to describe the emission and absorption of radiation, as well as the particulate properties of light itself, which by this time (in the wake of the discovery and interpretation of the Compton effect) was a generally accepted fact. On the other hand, the mechanics itself would have to be made relativistic, as it was known, already since the mid 1910s, that relativistic corrections to the kinematics of the electron would have a measurable effect in the fine structure of atomic spectra.

It was Schrödinger's wave mechanics, rather than matrix mechanics, that provided the ideal starting point for a relativistic kinematics of matter: One needed to find a new, relativistic matter wave equation, but at first glance there were no immediate other conceptual difficulties, such as the problem of a non-commuting time variable in matrix mechanics. Schrödinger himself had initially attempted to find a relativistic wave equation, following de Broglie's program of matter waves, which had been formulated in a relativistic manner. But de Broglie had stopped short of addressing the dynamical equations. Schrödinger in fact arrived at the Klein Gordon equation, but he dismissed it, due to its empirical inadequacy (later understood as the absence of spin in the Klein Gordon equation). Others were not as scrupulous, and the Klein Gordon equation was rediscovered (and published) multiple times in the immediate aftermath of Schrödinger's first papers on wave mechanics.[1]

Which role did general relativity play in this context? In full analogy with the case of special relativity, two distinct problems can be distinguished. On the one hand, there is the question of a quantum theory of the gravitational field. This will be discussed in the next part. On the other hand, there is the question of how quantum matter interacts with a gravitational field. We will be discussing this aspect first, because it is primary both logically and historically: Many of those who attempted a relativistic generalization of Schrödinger's program didn't see a reason to stop at special relativity—after all, the final quantum wave equation should also be compatible with general relativity and be able to describe the interaction of microscopic matter with a gravitational field. In two cases, the re-discovery of the Klein Gordon equation thus actually occurred in the context of curved space-time, in the work of Théophile de Donder and Oskar Klein.

[1]For a detailed history of the Klein Gordon equation, see Kragh (1984).

It should be noted that at the time, quantum mechanics was far from being well enough established for these works to be viewed as merely an application of the new quantum theory to general relativity. Rather, by incorporating wave mechanics within the better-established framework of general relativity, they sought to put the former on a surer foundation. From a modern viewpoint, De Donder's papers are notable for delivering the first construction of a Klein Gordon equation in curved spacetime (Chapter 7). But for de Donder this was not the central point. Rather, the generalization of Schrödingers "quantization" procedure for arriving at the Schrödinger equation to the generally relativistic case was in some sense meant to motivate the quantization procedure itself, even if de Donder's reasoning on this point is rather sketchy (Chapter 6).

In a similar vein, Klein's attempt to bring together wave mechanics with the five-dimensional extension of GR, proposed by Kaluza (1921) in order to unify gravity and electromagnetism, might simply be read as the construction of a Klein Gordon wave equation in curved spacetime with an electromagnetic potential. But Klein's main ambition was to see whether the interpretation of Schrödinger's waves as waves in actual space-time, which the likes of Bohr and Heisenberg doubted from the very start, might in fact be saved in a five-dimensional spacetime (Chapter 8). These hopes, cautiously voiced by Klein in the conclusion to his paper, were soon to be dashed, when Klein realized that such a theory would never be able to explain Planck's law of black body radiation, the problem that had led to quantum theory in the first place. In 1927, Klein converted to the probabilistic camp, now placing his hopes for a spatio-temporal description in the newly emerging quantum field theory.

This did not immediately change Klein's research program: He continued to elaborate his classical five-dimensional field theory, but now viewed it merely as a starting point for quantization. While work on a full theory of quantum electrodynamics was well underway at this time, the quantization of gravity was not viewed as an immediate concern. Klein's second paper (Chapter 9) thus begins with some remarks on why general relativity would need to be modified according to the postulates of quantum theory. This is the first published argument for the necessity of quantizing the gravitational field. He further argued that gravity and electrodynamics should be quantized simultaneously, that is, that one should quantize a unified field theory rather than simple Maxwellian electrodynamics (interacting with a Klein Gordon matter field), as this would, in particular, allow for a unification of the conservation laws of energy-momentum and of charge.

Indeed, if Klein's later recollections are to be believed, his five-dimensional field theory initially formed the classical basis for Heisenberg's and Pauli's attempts at formulating QED.[2] That is, until the Dirac equation came along in 1928 and ousted the Klein Gordon equation as the best description of relativistic quantum matter. Initially, Dirac saw one main advantage in his new equation over the old Klein Gordon one, namely that it could be interpreted as a one-particle, quantum mechanical Schrödinger equation (Dirac 1928). But even those who believed that any theory of matter waves would necessarily have to be second-quantized immediately realized the superiority of the Dirac equation, incorporating as it did the spin of the electron (and the proton).

In the wake of the Dirac equation, work immediately began on its integration into general relativity, the most immediate goal being the formulation of the Dirac equation in a curved (but non-dynamical) space-time. The first step in this direction was taken

[2] Oral history interview conducted by J. L. Heilbron and L. Rosenfeld on 25 February 1963, Niels Bohr Library and Archives, American Institute of Physics, College Park, MD, USA, http://www.aip.org/history/ohilist/4709_3.html.

by Hugo Tetrode, already in 1928 (Chapter 10). The general idea was the following: In Dirac's derivation, the flat Minkowski space-time metric $\eta_{\mu\nu}$ shows up in the algebra of the Dirac γ matrices:

$$\gamma_\mu\gamma_\nu + \gamma_\nu\gamma_\mu = 2\eta_{\mu\nu} \tag{5.1}$$

Tetrode's starting point was now to replace the Minkowski metric in this relation with a generalized non-flat metric $g_{\mu\nu}$, turning the (elements of the) γ matrices from constants into space-time dependent quantities, determined by the value of the metric at a given space-time point. But this determination was not unique, in fact it did not even offer a clear prescription for constructing the γ matrices when the metric was given. Tetrode only presented some tentative attempts at finding such a prescription, but, soon after, Eugene Wigner proposed that one could answer this question by making use of a formalism that Einstein had developed in that same year (1928), but in an entirely different context (Wigner 1929).

Einstein's idea was to replace the metric $g_{\mu\nu}$ as the fundamental quantity of his general theory of relativity with so-called "vierbeine" (tetrads), that is, for each point in space-time a local coordinate system spanned by four orthonormal basis vectors h_a^ν, with a running from one to four (Einstein 1928). The metric at a given point was determined by the tetrad to be

$$g_{\mu\nu} = h_{\mu a}\eta^{ab}h_{\nu b}, \tag{5.2}$$

but the metric by itself did not determine the tetrads. Indeed, the metric is unchanged by an arbitrary local Lorentz transformation of the tetrads

$$h_{\mu a} \to \Lambda_a^b h_{\mu b} \tag{5.3}$$

Einstein now proposed that the relative orientation of the tetrads was not redundant information, but actually carried physical meaning. In other words, once one has defined an x axis at some point in spacetime, it makes a difference what I call x axis in some other point—the only allowed transformations are then global Lorentz transformations, which conserve the relative orientation of the tetrads. This allowed one to speak of vectors at different spacetime points as parallel if they had the same components with regard to their local tetrad; hence Einstein's theory went by the name of teleparallelism. Through the notion of teleparallelism, the tetrads thus defined a second connection (beside the usual Levi-Civita connection one could obtain from the metric tensor alone) which could be used to construct new tensors and invariants, which could in turn be used to write down more general field equations. Einstein believed that he could thereby construct combined field equations for the gravitational and electromagnetic fields.

Wigner realized that, independent of these unified field theory considerations,[3] given local tetrads one could give a unique prescription for the construction of the γ matrices in curved spacetime as

[3]As we shall see, this is just one instance of the close connection between developments in unified field theory and the Dirac equation in curved space-time. For a detailed discussion from the viewpoint of the history of unified theory, see Goenner (2004).

$$\gamma_\mu = \tilde{\gamma}^a \eta_{ab} h^b_\mu \tag{5.4}$$

where the $\tilde{\gamma}$ are the usual (constant) Dirac matrices; the Dirac matrices on curved spacetime γ can then easily be shown to obey the correct anti-commutation relations. The open question was then: Given the Dirac matrices in curved spacetime, how were they to be incorporated into the Dirac equation? Both Tetrode and Wigner had tried this, but they had mainly concerned themselves with the question of how space-time dependent γ matrices were to be integrated into the Dirac equation, which of course contained derivatives with respect to space and time. The general covariance of the wave equations they ended up with was, however, very questionable. This question was soon taken up by Vladimir Fock, initially in collaboration with Dimitri Ivanenko.

Fock's main innovation was introducing the notion of the covariant derivative of a spinor, which allowed for a manifestly covariant Dirac equation in curved spacetime (Chapter 11). The connection appearing in the covariant spinor derivative was determined by the Dirac matrices in curved spacetime, which in turn were determined by the tetrads. But Fock could show that his Dirac equation was indeed covariant under local Lorentz transformations, that is, that the choice of tetrads (beyond what was determined by the metric) played no physical role. The tetrads were thus demoted to mere mathematical tools, in contrast with the physical role they had played in Einstein's teleparallelism.

A very similar approach was developed independently by Hermann Weyl (Chapter 12).[4] But Weyl's approach differed fundamentally from Fock's in its perceived goal. Fock came from quantum theory and viewed his work as an extension of Dirac's theory to the case of curved spacetime. Weyl, on the other hand, is a representative of the viewpoint, which we already encountered in the work of de Donder and the early Klein, that general relativity might help in solving foundational issues in quantum mechanics. For while the original problems of motivating the quantization procedure or interpreting the wave function had largely been set aside by 1929 (the year of Weyl's and Fock's work), the Dirac equation brought with it new fundamental difficulties, most notably the problem of the negative energy states.

Weyl attempted to solve this difficulty by replacing the four-spinors of Dirac with the two-spinors that now carry Weyl's name. This removed the negative energy solutions, but at the same time also prevented the inclusion of the usual explicit mass term in the wave equation. Weyl expressed the rather vague hope that the mass might arise through the coupling to a gravitational field and, to this end, constructed an essentially classical field theory of coupled spinor, gravitational and electromagnetic fields, which, just like Klein's five-dimensional field theory two years earlier, was eventually supposed to be quantized. In order to couple the spinors to the gravitational field, he introduced the same tetrad-based covariant spinor derivative that Fock had constructed. For the coupling of the spinors to the electromagnetic field, he discovered, almost in passing, the local gauge invariance of the Dirac equation, which he not only related to, but in fact tried to deduce from local Lorentz invariance.

The work of Weyl and Fock was received in two distinct ways: Some, coming to the problem from general relativity, attempted to further pursue the solution of the negative energy problem along these lines, most notably Jan Schouten, who attempted to show how a mass term for Weyl spinors might arise in five- instead of four-dimensional curved spacetime (Schouten 1931). In the quantum community, on the other hand, the work of

[4]For a detailed discussion of Weyl's, and also Fock's, work and its context, see Scholz (2005).

Weyl and Fock was viewed as the classical basis for a relativistic quantum field theory of gravitation, electromagnetism and charged Dirac particles, and an attempt to quantize it was undertaken almost immediately by Léon Rosenfeld (see the following part of this book). It was important for this undertaking that Weyl and Fock had not only shown how to couple spinor waves to a gravitational field, but had also identified the degrees of freedom in general relativity that were supposed to be subjected to quantization in such a coupled theory, the tetrads.

In 1932, Schrödinger showed that this was not the unique way of integrating spinors into general relativity. He took the spacetime dependent Dirac matrices as fundamental degrees of freedom, instead of the tetrads (Chapter 13). This brought the question of which basis to choose for the Dirac matrices back onto the table: In the tetrad prescription, equation 5.4, one simply had to initially pick some basis for the constant Dirac matrices; everything else then followed from the tetrads and their transformation properties. But if the γ matrices themselves were supposed to be the dynamical variables, no direct reference could be made to the constant matrices. Schrödinger thus imposed further hermiticity conditions, which defined the space-time dependent Dirac matrices, given a metric, up to a unitary transformation in spin space. These unitary transformations were not only, as Schrödinger put it, "benign and irrelevant," they also corresponded exactly to the local Lorentz transformations in the Weyl-Fock tetrad formalism. What had thus, using tetrads, looked like the rotation of local coordinate axes, in Schrödinger's formalism was interpreted as a (local) transformation in spin space. This latter point, in particular, was soon after further worked out by Bargmann (1932) and Pauli (1933).

A further alternative was worked out by Bartel van der Waerden and Leopold Infeld (Chapter 14), based on a generalization of van der Waerden's special relativistic spinor calculus (Waerden 1929).[5] They defined, instead of the local coordinate basis of the tetrad formalism, local spinor spaces. The separation between local spin space transformations and coordinate transformations was thus now taken as the starting point, rather than as a result of the analysis. As in special relativistic spinor calculus, where vectors are constructed from spinors through the Pauli matrices, the connection between spinor space and spacetime was established through a set of local, spacetime dependent generalized Pauli matrices. These matrices replaced the metric as fundamental degrees of freedom. They were thus the two-spinor analog of Schrödinger's spacetime dependent Dirac matrices. In a sense, Infeld/van der Waerden was to Schrödinger as Weyl was to Fock: The same general idea, but with two-spinors instead of four-spinors. Van der Waerden viewed his formalism as a considerable simplification of Schrödinger's four-dimensional framework, since one no longer needed to employ hermiticity conditions in order to fix the dynamical matrix degrees of freedom.[6]

Both the Schrödinger and the Infeld-van der Waerden formalism had to wait quite a while until they were put to use. Even though Pauli immediately suggested to Rosenfeld that he should repeat his analysis of 1930, this time quantizing the Dirac matrices instead of tetrads,[7] Schrödinger's reformulation was, to the best of my knowledge, not used as the basis for a quantization of gravity until Bryce de Witt's thesis work in the late 1940s. And, still in 1951,[8] Schrödinger remarked to Fredrik Belinfante: "I do not set any store in my paper of 1932 on Dirac's theory in a general metric. It was at the time a smart exercise,

[5] For a detailed study of van der Waerden's work in physics, including this paper, see Schneider (2010).
[6] See the letter from van der Waerden to Schrödinger, 14 June 1932, Schrödinger papers, Vienna University Library.
[7] Letter from Pauli to Rosenfeld, 25 November 1932, reprinted in Meyenn (1993).
[8] Letter from Schrödinger to Belinfante, 16 March 1951, Archive for the History of Quantum Physics.

but it came to nothing." Van der Waerden was equally dismissive of his work. His formalism had to wait even longer, until the late 1950s, before it was widely received, mainly, however, in the context of mathematical relativity, rather than as a basis for quantization.[9]

In any case, by the early 1930s it appeared that the classical field theory of relativistic quantum matter interacting with the gravitational field had been worked out, in several closely related mathematical formulations. The core difficulties were seen to lie in the quantization procedure, which, already for the simpler case of electrodynamic interactions, led to uninterpretable infinities in the higher orders of perturbation theory.

In 1938, however, a study by Einstein, Infeld, and Banesh Hoffmann (EIH) appeared, which appeared to indicate that there were serious advantages to starting with a classical theory of matter as point particles, in contradiction, ironically, with Einstein's program of a unified field theory (Chapter 15). EIH could show that if one treated matter particles as singularities in the gravitational field , one could (in stepwise approximations) derive the equations of motion for these singularities from the field equations alone, without having to additionally postulate the geodesic equations for particle motion. The prospect of carrying this feature of general relativity over to the quantum theory made a classical theory of point particles (rather than a classical theory of matter waves to be second-quantized) an attractive starting point for a quantum theory of gravitation. This approach was taken by Peter Bergmann, as will be discussed in the final part of this book.

Several different possible starting points for a quantization of gravity, that is, a modification of general relativity according to the principles of quantum theory, had thus been set. But the inverse program, the modification of quantum theory following considerations from general relativity, was not dead. As we have seen, this program by the late 1920s mainly centered around the idea that the inclusion of general relativity might somehow determine the mass of the elementary particles, which were after all purely empirical parameters in the wave equations of quantum mechanics and field theory.

A variant of this idea, which brought in ideas from relativistic cosmology, had been formulated early on by Arthur Eddington (1923). He hypothesized that the radius of the electron (proportional to the inverse mass) should be related (in fact proportional) to the curvature radius R of the universe as a whole, still viewed at the time as static and spherical (an Einstein universe). Eddington further elaborated on this idea 13 years later in the context of his notorious and oft-ridiculed relativistic theory of protons and electrons (Eddington 1936),[10] obtaining the necessarily huge proportionality constant through the square root of the number N of electrons (or protons) in the entire universe.

While this work was almost universally regarded as obscure and incomprehensible, some of the general ideas concerning the relation between cosmology and elementary particles were taken up by Erwin Schrödinger.[11] Schrödinger saw some promise in Eddington's new work, inasmuch as it was a specification of the original idea of connecting the electron mass with cosmology, now also incorporating the tools of wave mechanics. In particular, Eddington had derived the rest mass energy of an elementary particle as the energy of the first excited state for the system of all the identical fermions in the universe. The curvature radius of the universe entered through the boundary conditions for the wave equation, the total number of particles in the universe through the fact that one was looking at energies near the Fermi surface. In a stripped-down version, which Schrödinger

[9] See, in particular, Penrose (1960).

[10] For a benevolent and insightful reading of this book, see Kilmister (1994). For more on the historical background, see Kragh (2015).

[11] For more on this work in the context of Schrödinger's work on cosmology more generally, see Urbantke (1992).

presented at the Galvani Bicentennial in 1937 (Schrödinger 1938b) and which made no more reference to the more eccentric aspects of Eddington's theory (such as the attempt to prove that the fine structure constant was necessarily the reciprocal of an integer),[12] the central relation read:

$$mc^2 = \cdots \frac{hc\sqrt{N}}{R}, \tag{5.5}$$

where the dots refer to undetermined constants. In order to obtain such a relation from the accepted wave mechanics, rather than from Eddington's novel theory, Schrödinger devoted considerable energy to the explicit construction of the solutions of the Maxwell and Dirac equations in a spherical universe (Schrödinger 1938a), also studying the non-static, expanding case, which was by that time heavily favored (Schrödinger 1940). But after three years of working on this subject, he had to conclude in a small note that one would, starting from very general assumptions, always obtain the cubic root of N in Eddington's relation, instead of the square root, and thus obtain much too small masses for the elementary particles (Chapter 16).

The idea that general relativity, in particular relativistic cosmology, might directly determine the microscopic quantum wave equations thus remained nothing more than a tantalizing possibility. In contrast, the program of quantizing the gravitational field, while technically challenging, seemed a fairly straightforward task. In this section, we have discussed the necessary groundwork for this task, that is, the establishment of classical theories of matter interacting with the gravitational field that went beyond the very unspecific energy-momentum tensor of matter appearing in the Einstein equations. In the next section, we will discuss the first attempts at an actual quantization of such theories.

References

Bargmann, Valentine (1932). Bemerkungen zur allgemein-relativistischen Fassung der Quantentheorie. *Sitzungsberichte der Preussischen Akademie der Wissenschaften*: 346–354.

Dirac, Paul A. M. (1928). Zur Quantentheorie des Elektrons. In: *Quantentheorie und Chemie*. Ed. by Hans Falkenhagen. Leipzig: S. Hirzel, 85–94.

Eddington, Arthur S. (1923). *The Mathematical Theory of Relativity*. Cambridge: Cambridge University Press.

— (1936). *Relativity Theory of Protons and Electrons*. London: Cambridge University Press.

Einstein, Albert (1928). Riemann-Geometrie mit Aufrechterhaltung des Begriffes des Fernparallelismus. *Sitzungsberichte der Preussischen Akademie der Wissenschaften*: 217–221.

Goenner, Hubert F. M. (2004). On the History of Unified Field Theories. *Living Reviews in Relativity* 7(2):1–153.

Kaluza, Theodor (1921). Zum Unitätsproblem der Physik. *Sitzungsberichte der Preussischen Akademie der Wissenschaften*:966–972.

[12]Still, reactions were mixed. On 18 February 1938, Sommerfeld wrote to Schrödinger: "It is wonderful that you are now seriously working on Eddingt.[on]." (*Es ist wundervoll, dass Sie mit Eddingt.[on] Ernst machen*) but also talked about Bohr's "arrogant dismissal" (*bonzenhafte Ablehnung*), Schrödinger papers, Vienna University Library.

Kilmister, Clive W. (1994). *Eddington's Search for a Fundamental Theory: A Key to the Universe*. Cambridge: Cambridge University Press.

Kragh, Helge (1984). Equation with the Many Fathers: The Klein-Gordon Equation in 1926. *American Journal of Physics* 52(11):1024–1033.

— (2015). On Arthur Eddington's Theory of Everything. *arXiv:1510.04046*.

Meyenn, Karl von, ed. (1993). *Wolfgang Pauli: Wissenschaftlicher Briefwechsel mit Bohr, Einstein, Heisenberg u.a.* Vol. 3: 1940–1949. Berlin: Springer.

Pauli, Wolfgang (1933). Über die Formulierung der Naturgesetze mit fünf homogenen Koordinaten. II. Die Diracschen Gleichungen für die Materiewellen. *Annalen der Physik* 5(18):337–372.

Penrose, Roger (1960). A Spinor Approach to General Relativity. *Annals of Physics* 10(2): 171–201.

Schneider, Martina R. (2010). *Die physikalischen Arbeiten des jungen B. L. van der Waerden*. PhD thesis. Bergische Universität Wuppertal.

Scholz, Erhard (2005). Local Spinor Structures in V. Fock's and H. Weyl's Work on the Dirac Equation (1929). In: *Géométrie au vigntième siècle, 1930–2000*. Ed. by D. Flament, J. Kouneiher, P. Nabonnand, and J.-J. Szsceciniarz. Paris: Hermann, 284–301.

Schouten, Jan A. (1931). Dirac Equations in General Relativity. *Journal of Mathematics and Physics* 10:239–283.

Schrödinger, Erwin (1938a). Eigenschwingungen des sphärischen Raumes. *Commentationes Pontificiae Academiae Scientiarum* 2:321–364.

— (1938b). Sur la théorie du monde d'Eddington. *Il Nuovo Cimento* 15:246–254.

— (1940). Maxwell's and Dirac's Equations in the Expanding Universe. *Proceedings of the Royal Irish Academy* 46A:25–47.

Urbantke, Helmuth (1992). Schrödinger and Cosmology. In: *Studies in the History of General Relativity, Einstein Studies Vol. 3*. Ed. by Jean Eisenstaedt and A. J. Kox. Boston: Birkhäuser, 453–459.

Waerden, Bartel van der (1929). Spinoranalyse. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Math.-Phys. Klasse*:100–109.

Wigner, Eugene P. (1929). Eine Bemerkung zu Einsteins neuer Formulierung des allgemeinen Relativitätsprinzips. *Zeitschrift für Physik* 53(7):592–596.

Chapter 6
Théophile de Donder and Frans Henri van den Dungen (1926): La quantification déduite de la Gravifique einsteinienne

Théophile de Donder and Frans Henri van den Dungen (1926). La quantification déduite de la Gravifique einsteinienne. *Comptes rendus*, 183: 22–24.

abscisses commune à nos deux travaux, ne donne pas un très bon accord : la courbe de Relf, moins rapidement ascendante, coupe la mienne vers $R = 350$, sous un angle notable.

En résumé, mes expériences sur les tourbillons alternés dans les liquides conduisent à des résultats en désaccord appréciable avec la loi de similitude dynamique, écarts qui me paraissent ne pouvoir être tous attribués à des erreurs de mesures [1].

PHYSIQUE MATHÉMATIQUE. — *La quantification déduite de la Gravifique einsteinienne.* Note de MM. **Th. De Donder** et **Fr. H. van den Dungen**, présentée par M. Hadamard.

Reportons-nous à l'équation (9) de la Note [2] parue dans ces *Comptes rendus* et due à l'un des auteurs de celle-ci. Introduisons la fonction invariante F_ν au moyen de

$$kS_\nu = \ln F_\nu \qquad (\nu = 1, \ldots, N) \tag{1}$$

où k représente une constante universelle. Cette équation (9) devient alors

$$I_\nu \equiv \sum_\varphi \sum_\psi g_*^{\varphi\psi} \left(\frac{\partial F_\nu}{\partial q_\varphi} - k\tau_\nu^{(e)} \Phi_\varphi^* F_\nu \right) \left(\frac{\partial F_\nu}{\partial q_\psi} - k\tau_\nu^{(e)} \Phi_\psi^* F_\nu \right) - (k\tau_\nu^{(m)} F_\nu)^2 = 0. \tag{2}$$

Prenons, par rapport à F_ν, la dérivée variationnelle du multiplicateur $I\sqrt{-g_*}$; d'où

$$\frac{\delta(I_\nu\sqrt{-g_*})}{\delta F_\nu} \equiv \frac{\partial(I_\nu\sqrt{-g_*})}{\partial F_\nu} - \sum_{\varphi=1}^{f} \frac{\partial}{\partial q_\varphi} \left[\frac{\partial(I_\nu\sqrt{-g_*})}{\partial \frac{\partial F_\nu}{\partial q_\varphi}} \right]. \tag{3}$$

---

(1) C'est aussi la conclusion de E. G. Richardson, auteur d'un autre travail récent sur les sons éoliens (*Proc. Phys. Soc. of London*, 36, 1924, p. 153); les fréquences des tourbillons y sont mesurées par les méthodes les plus variées dans l'air, dans l'eau et dans une solution de mélasse, R variant de 32 à 543, et S de 0,09 à 0,30, donc à peu près comme dans mes propres expériences.

Lors de la discussion qui a suivi la Communication de E. G. Richardson à la *Physical Society*, C. V. Drysdale propose comme une idée nouvelle ma méthode optique des rayons aberrants pour enregistrer les tourbillons.

(2) Th. De Donder, *Application de la relativité aux systèmes atomiques et moléculaires* (*Comptes rendus*, 182, 1926, p. 1380-1382).

En développant les calculs dans (3), on est amené à poser :

$$\Delta_q F_\nu \equiv \frac{1}{\sqrt{-g_*}} \sum_\varphi \sum_\psi \frac{\partial}{\partial q_\varphi}\left(g_*^{\varphi\psi}\sqrt{-g_*}\frac{\partial F_\nu}{\partial q_\psi}\right); \tag{4}$$

$$D \equiv \frac{1}{\sqrt{-g_*}} \sum_\varphi \sum_\psi \frac{\partial}{\partial q_\varphi}(g_*^{\varphi\psi}\sqrt{-g_*}\,\Phi_\psi^*); \tag{5}$$

$$E \equiv \sum_\varphi \sum_\psi g_*^{\varphi\psi}\Phi_\varphi^*\Phi_\psi^*. \tag{6}$$

Égalant la dérivée variationnelle (3) *à zéro*, c'est-à-dire *en extrémant*

$$\int I_\nu\sqrt{-g_*}\,dq_1\ldots dq_f,$$

nous obtenons, en utilisant les notations (4), (5) et (6),

$$\Delta_q F_\nu - k\tau_\nu^{(e)} D F_\nu - k^2(E\tau_\nu^{(e)^2} - \tau_\nu^{(m)^2})F_\nu = 0. \tag{7}$$

Nous pourrons disposer des potentiels Φ_ψ^* ($\psi = 1, \ldots, f$) de manière qu'on ait $D = 0$; c'est la généralisation de l'équation complémentaire de Maxwell. Ainsi, (7) devient

$$\Delta_q F_\nu + R_\nu F_\nu = 0 \qquad (\nu = 1, \ldots, N), \tag{8}$$

où l'on a écrit

$$R_\nu \equiv k^2(\tau_\nu^{(m)^2} - E\tau_\nu^{(e)^2}). \tag{9}$$

L'équation (8) doit être vérifiée en tout point de l'espace des configurations $q_1, \ldots, q_f$; elle peut être remplacée par l'*équation intégrale de Fredholm :*

$$F_{\nu P} = \int_q G_{\nu PM} R_{\nu M} F_{\nu M}\,\delta q_1\ldots\delta q_f,$$

P et M étant deux points de l'espace q et $G_{\nu PM}$ représentant la fonction de Green correspondante, à un facteur constant près. On a supposé que F_ν est donné sur la surface limitant l'espace q, ou, si l'espace q est illimité, que F_ν tend vers zéro comme $r^{(1-f)}$ pour $f > 1$, ou comme $\log r$, pour $f = 1$; le symbole r représente la distance radiale dans l'espace des configurations défini par

$$(\delta s)^2 \equiv \sum_\varphi \sum_\psi g_{\varphi\psi}^*\,\delta q_\varphi\,\delta q_\psi.$$

L'équation intégrale (10) peut encore être écrite quand le phénomène est périodique par rapport à certains degrés de liberté.

L'équation intégrale homogène (10) n'admettra de solution F_ν non identiquement nulle que pour une suite de valeurs déterminées et distinctes (¹) de R_ν. On obtient ainsi *la quantification grâce à la Gravifique einsteinienne.*

Ces résultats sont à rapprocher des importantes recherches de M. E. Schrödinger (²) concernant la quantification de certains systèmes *non* relativistiques.

PHYSIQUE MATHÉMATIQUE. — *La mécanique ondulatoire de Schrödinger; une méthode générale de résolution par approximations successives.* Note de M. **Léon Brillouin**, présentée par M. Marcel Brillouin.

1. Schrödinger, développant les idées de L. de Broglie, a précisé récemment les grands traits d'une mécanique atomique ondulatoire (³). Soit un système atomique, dont l'énergie potentielle est $V(q^1, \ldots, q^n)$, tandis que l'énergie cinétique T a pour expression

$$(1) \qquad 2T = \sum_{k,l} m_{kl}\,\dot{q}^k\,\dot{q}^l = \sum_{k,l} m^{kl}\,p_k\,p_l.$$

Les m^{kl} sont des fonctions des coordonnées $q^k\ldots$, et nous appellerons m le déterminant des m^{kl}. L'équation classique de Hamilton s'écrit

$$(2) \qquad \sum_{k,l} m^{kl}\,\frac{\partial W_0}{\partial q^k}\,\frac{\partial W_0}{\partial q^l} + 2(V - E) = 0,$$

E représentant la constante de l'énergie. Schrödinger aboutit à l'équation générale suivante :

$$(3) \qquad m^{\frac{1}{2}} \sum_{k,l} \frac{\partial}{\partial q^k}\left\{ m^{-\frac{1}{2}}\, m^{kl}\,\frac{\partial\psi}{\partial q^l} \right\} - \frac{8\pi^2}{h^2}(V - E)\psi = 0,$$

où h est la constante de Planck. Les niveaux d'énergie quantifiés E sont les

---

(¹) Nous supposons ici que le noyau de l'équation (10) ne possède pas de points d'indétermination ou de singularités essentielles.

(²) *Annalen der Physik*, Leipzig, 1926, vierte Folge, 79, p. 361-376, 489-527, 734-756.

(³) L. de Broglie, *Thèse;* E. Schrödinger, *Ann. der Phys.*, 79, 1926, p. 361, 489, 734; l'équation (3) ci-dessus se trouve dans Schrödinger, p. 310, équation (18), et p. 748, équation (31).

Chapter 7
Théophile de Donder (1926): Application de la quantification déduite de la Gravifique einsteinienne

Théophile de Donder (1926). Application de la quantification déduite de la Gravifique einsteinienne. *Comptes rendus*, 183: 594–595.

Les considérations ci-dessus montrent que, pour étudier le rendement économique d'un carburant, la connaissance du pouvoir calorifique est absolument insuffisante et qu'il faut introduire la notion du rendement de combustion dans lequel interviennent à la fois la constitution physique du mélange et ses propriétés chimiques vis-à-vis des phénomènes limitateurs de compression.

PHYSIQUE MATHÉMATIQUE. — *Application de la quantification déduite de la Gravifique einsteinienne.* Note de M. **Th. De Donder**, présentée par M. Hadamard.

Dans une Note précédente [1] nous avons montré, dans l'espace-temps, comment on pouvait déduire la quantification des systèmes moléculaires de la Gravifique einsteinienne. Nous nous proposons maintenant d'appliquer cette méthode à la quantification d'un électron se mouvant dans un champ gravifique et électromagnétique quelconque.

Reportons-nous à l'équation (360) de la *Gravifique einsteinienne* [2], à savoir

$$(1) \qquad -H = c^2 m' V^{-1} \sum_{a=1}^{4} g_{a4} v^a + e' \Phi_4,$$

où e' et m' sont la charge et la masse de l'électron au repos. Supposons que la fonction hamiltonienne H soit indépendante du temps t; alors, H est un invariant du mouvement de l'électron considéré. Posons

$$(2) \qquad -H \equiv c(E + c^2 m');$$

en vertu de (1), nous pouvons dire que E est l'énergie totale de l'électron en mouvement, non compris son contenu énergétique $c^2 m'$.

Considérons aussi l'équation (362) de la *Gravifique* citée ci-dessus; à savoir :

$$(3) \qquad \sum_\alpha \sum_\beta g^{\alpha\beta}(p_\alpha - e'\Phi_\alpha)(p_\beta - e'\Phi_\beta) - (c^2 m')^2 = 0 \qquad (\alpha, \beta = 1, 2, 3, 4).$$

Remplaçons p_4 par sa valeur $-H$, c'est-à-dire $c(E + c^2 m')$; d'où l'équa-

---

(1) Th. De Donder et Fr. H. van den Dungen, *Comptes rendus*, 182, 1926, p. 22-24.

(2) Th. De Donder, *La Gravifique einsteinienne*, Paris, Gauthier-Villars, 1921, p. 85.

tion de Jacobi, en posant $k\mathrm{S} = ln\Psi$,

$$
\begin{aligned}
(4)\qquad \mathrm{I} \equiv & \sum_i \sum_j g^{ij}\left(\frac{\partial\Psi}{\partial x_i} - e'\Phi_i k\Psi\right)\left(\frac{\partial\Psi}{\partial x_j} - e'\Phi_j h\Psi\right) \\
& + 2k\Psi \sum_i g^{4i}(c\mathrm{E} + c^3 m' - e'\Phi_4)\left(\frac{\partial\Psi}{\partial x_i} - e'\Phi_i k\Psi\right) \\
& + k^2\Psi^2[g^{44}(c\mathrm{E} + c^3 m' - e'\Phi_4)^2 - (c^2 m')^2] = 0 \\
& \qquad (i, j = 1, 2, 3).
\end{aligned}
$$

Comme nous l'avons fait dans notre Note citée ci-dessus, annulons la dérivée variationnelle de $\mathrm{I}\sqrt{-g}$ par rapport à Ψ; d'où

$$
\begin{aligned}
(5)\quad \Delta\Psi - k\Psi \Big[& e'\mathcal{D} + k\mathcal{E}e'^2 - k(c^2 m')^2 + kg^{44}(c\mathrm{E} + c^3 m' - e'\Phi_4)^2 \\
& - 2k\sum g^{4i}(c\mathrm{E} + c^3 m' - e'\Phi_4)e'\Phi^i \\
& + \sum_i \frac{\partial}{\partial x_i}[g^{4i}(c\mathrm{E} + c^3 m' - e'\Phi_4)]\Big] = 0.
\end{aligned}
$$

où nous avons posé

$$
(6)\qquad \left\{
\begin{aligned}
\Delta\Psi &\equiv \frac{1}{\sqrt{-g}} \sum_i \sum_j \frac{\partial}{\partial x_i}\left(g^{ij}\sqrt{-g}\,\frac{\partial\Psi}{\partial x_i}\right) \\
\mathcal{D} &\equiv \frac{1}{\sqrt{-g}} \sum_i \sum_j \frac{\partial}{\partial x_i}(g^{ij}\sqrt{-g}\,\Phi_j) \\
\mathcal{E} &\equiv \sum_i \sum_j g^{ij}\Phi_i\Phi_j \qquad (i, j = 1, 2, 3)
\end{aligned}
\right.
$$

L'équation aux dérivées partielles du second ordre en Ψ fournira la quantification de l'électron considéré.

Comme *exemple*, prenons *un champ de Minkowski* dans lequel l'électron se meut ; l'équation (5) devient :

$$
(7)\quad \Delta\Psi - k\Psi\left[e'\left(ke'\mathcal{E}\frac{1}{c^2}\frac{\partial\Phi_4}{\partial t} + k2m'e^2\left(\mathrm{E} - \frac{e'\Phi^4}{c}\right) + k\left(\mathrm{E} - \frac{e'\Phi_4}{c}\right)^2\right]\right. = 0.
$$

Si nous supposons enfin que $\Phi_1 = \Phi_2 = \Phi_3 = 0$, on aura, *en première approximation*

$$
(8)\qquad \Delta\Psi - k^2\Psi\, 2m'c^2(\mathrm{E} - \mathrm{V}) = 0
$$

où l'on a posé $\mathrm{V} \equiv \dfrac{e'\Phi_4}{c}$. L'équation (8) est celle utilisée par M. E. Schrödinger (¹) dans sa mécanique ondulatoire.

(¹) E. SCHRÖDINGER, *Annalen der Physik*, 79, 1926, p. 361-376. Voir spécialement équations (5), (23) et (24).

Chapter 8
Oskar Klein (1926): Quantentheorie und fünfdimensionale Relativitätstheorie

Oskar Klein (1926). Quantentheorie und fünfdimensionale Relativitätstheorie. *Zeitschrift für Physik* 37: 895–906.

Quantentheorie und fünfdimensionale Relativitätstheorie.

Von **Oskar Klein** in Kopenhagen.

(Eingegangen am 28. April 1926.)

Auf den folgenden Seiten möchte ich auf einen einfachen Zusammenhang hinweisen zwischen der von Kaluza [1]) vorgeschlagenen Theorie für den Zusammenhang zwischen Elektromagnetismus und Gravitation einerseits und der von de Broglie [2]) und Schrödinger [3]) angegebenen Methode zur Behandlung der Quantenprobleme andererseits. Die Theorie von Kaluza geht darauf hinaus, die zehn Einsteinschen Gravitationspotentiale g_{ik} und die vier elektromagnetischen Potentiale φ_i in Zusammenhang zu bringen mit den Koeffizienten γ_{ik} eines Linienelementes von einem Riemannschen Raum, der außer den vier gewöhnlichen Dimensionen noch eine fünfte Dimension enthält. Die Bewegungsgleichungen der elektrischen Teilchen nehmen hierbei auch in elektromagnetischen Feldern die Gestalt von Gleichungen geodätischer Linien an. Wenn dieselben als Strahlengleichungen gedeutet werden, indem die Materie als eine Art Wellenausbreitung betrachtet wird, kommt man fast von selbst zu einer partiellen Differentialgleichung zweiter Ordnung, die als eine Verallgemeinerung der gewöhnlichen Wellengleichung angesehen werden kann. Werden nun solche Lösungen dieser Gleichung betrachtet, bei denen die fünfte Dimension rein harmonisch auftritt mit einer bestimmten mit der Planckschen Konstante zusammenhängenden Periode, so kommt man eben zu den obenerwähnten quantentheoretischen Methoden.

§ 1. Fünfdimensionale Relativitätstheorie. Ich fange damit an, eine kurze Darstellung von der fünfdimensionalen Relativitätstheorie zu geben, die sich nahe an die Theorie von Kaluza anschließt, aber in einigen Punkten von derselben abweicht.

Betrachten wir ein fünfdimensionales Riemannsches Linienelement, für welches wir einen vom Koordinatensystem unabhängigen Sinn postulieren. Wir schreiben dasselbe:

$$d\sigma = \sqrt{\Sigma \gamma_{ik}\, d x^i\, d x^k}, \tag{1}$$

wo das Zeichen Σ, wie überall im folgenden, eine Summation über die doppelt vorkommenden Indizes von 0 bis 4 angibt. Hierbei bezeichnen $x^0 \ldots x^4$ die fünf Koordinaten des Raumes. Die 15 Größen γ_{ik} sind die kovarianten Komponenten eines fünfdimensionalen symmetrischen Tensors. Um von denselben zu den Größen g_{ik} und φ_i der gewöhnlichen Relativitätstheorie zu kommen, müssen wir gewisse spezielle Annahmen machen. Erstens müssen vier der Koordinaten, sagen wir x^1, x^2, x^3, x^4, stets den gewöhnlichen Zeitraum charakterisieren. Zweitens dürfen die Größen

1) Th. Kaluza, Sitzungsber. d. Berl. Akad. 1921, S. 966.

2) L. de Broglie, Ann. d. Phys. (10) **3**, 22, 1925. Thèses, Paris 1924.

3) E. Schrödinger, Ann. d. Phys. **79**, 361 und 489, 1926.

γ_{ik} nicht von der fünften Koordinate x^0 abhängen. Hieraus folgt, daß die erlaubten Koordinatentransformationen sich auf die folgende Gruppe beschränken [1]):

$$\left.\begin{aligned} x^0 &= x^{0\prime} + \psi_0(x^{1\prime}, x^{2\prime}, x^{3\prime}, x^{4\prime}), \\ x^i &= \psi_i(x^{1\prime}, x^{2\prime}, x^{3\prime}, x^{4\prime}) \qquad (i = 1, 2, 3, 4). \end{aligned}\right\} \tag{2}$$

Eigentlich hätten wir in der ersten Gleichung Konstante mal $x^{0\prime}$ anstatt $x^{0\prime}$ schreiben sollen. Die Beschränkung auf den Wert Eins der Konstante ist ja aber ganz unwesentlich.

Wie man leicht zeigt, bleibt γ_{00} bei den Transformationen (2) invariant. Die Annahme $\gamma_{00} =$ Const. ist deshalb zulässig. Die Vermutung liegt nahe, daß nur die Verhältnisse der γ_{ik} einen physikalischen Sinn haben. Dann ist diese Annahme nur eine immer mögliche Konvention. Indem wir die Maßeinheit von x^0 vorläufig unbestimmt lassen, setzen wir:

$$\gamma_{00} = \alpha. \tag{3}$$

Man zeigt ferner, daß die folgenden Differentialgrößen bei den Transformationen (2) invariant bleiben, nämlich [1]):

$$d\vartheta = dx^0 + \frac{\gamma_{0i}}{\gamma_{00}}\, dx^i, \tag{4}$$

$$ds^2 = \left(\gamma_{ik} - \frac{\gamma_{0i}\gamma_{0k}}{\gamma_{00}}\right) dx^i\, dx^k. \tag{5}$$

In diesen Ausdrücken soll über die doppelt vorkommenden Indizes von 1 bis 4 summiert werden. Bei solchen Summen wollen wir, wie üblich, das Summenzeichen fortlassen. Die Größen $d\vartheta$ und ds hängen in der folgenden Weise mit dem Linienelement $d\sigma$ zusammen:

$$d\sigma^2 = \alpha\, d\vartheta^2 + ds^2. \tag{6}$$

Auf Grund der Invarianz von $d\vartheta$ und γ_{00} folgt nun, daß die vier γ_{0i} ($i \neq 0$), wenn x^0 festgehalten wird, sich wie die kovarianten Komponenten eines gewöhnlichen Vierervektors transformieren. Wenn x^0 mittransformiert wird, tritt noch der Gradient eines Skalars additiv hinzu. Dies bedeutet, daß die Größen:

$$\frac{\partial \gamma_{0i}}{\partial x^k} - \frac{\partial \gamma_{0k}}{\partial x^i}$$

[1]) Vgl. H. A. Kramers, Proc. Amsterdam **23**, Nr. 7, 1922, wo eine an die nun folgenden Betrachtungen erinnernde Überlegung mit einem einfachen Beweis für die Invarianz von $d\vartheta$ und ds^2 gegeben ist.

sich wie die kovarianten Komponenten F_{ik} des elektromagnetischen Feldtensors transformieren. Die Größen γ_{0i} verhalten sich also vom invariantentheoretischen Gesichtspunkt wie die elektromagnetischen Potentiale φ_i. Wir nehmen deshalb an:

$$d\vartheta = dx^0 + \beta\varphi_i dx^i, \tag{7}$$

d. h.

$$\gamma_{0i} = \alpha\beta\varphi_i \qquad (i = 1, 2, 3, 4), \tag{8}$$

wo β eine Konstante bedeutet, und wo die φ_i so definiert sind, daß in rechtwinkligen Galileischen Koordinaten gilt:

$$\left.\begin{aligned} (\varphi_x, \varphi_y, \varphi_z) &= A, \\ \varphi_z &= -cV, \end{aligned}\right\} \tag{9}$$

wo A das gewöhnliche Vektorpotential, V das gewöhnliche skalare Potential und c die Lichtgeschwindigkeit bezeichnen.

Die Differentialform ds wollen wir mit dem Linienelement der gewöhnlichen Relativitätstheorie identifizieren. Wir setzen also

$$\gamma_{ik} = g_{ik} + \alpha\beta^2\varphi_i\varphi_k, \tag{10}$$

wobei wir die g_{ik} so wählen wollen, daß in rechtwinkligen Galileischen Koordinaten gilt:

$$ds^2 = dx^2 + dy^2 + dz^2 - c^2 dt^2. \tag{11}$$

Hiermit sind die Größen γ_{ik} auf bekannte Größen zurückgeführt. Das Problem ist nun, solche Feldgleichungen für die Größen γ_{ik} aufzustellen, daß sich für die g_{ik} und φ_i in genügender Annäherung die Feldgleichungen der gewöhnlichen Relativitätstheorie ergeben. Auf dieses schwierige Problem wollen wir hier nicht näher eingehen, sondern wir wollen nur zeigen, daß die gewöhnlichen Feldgleichungen von dem Gesichtspunkt der fünfdimensionalen Geometrie sich einfach zusammenfassen lassen. Wir bilden die Invariante:

$$P = \sum \gamma^{ik}\left[\frac{\partial \begin{Bmatrix} i\,\mu \\ \mu \end{Bmatrix}}{\partial x^k} - \frac{\partial \begin{Bmatrix} i\,k \\ \mu \end{Bmatrix}}{\partial x^\mu} + \begin{Bmatrix} i\,\mu \\ \nu \end{Bmatrix}\begin{Bmatrix} k\,\nu \\ \mu \end{Bmatrix} - \begin{Bmatrix} i\,k \\ \mu \end{Bmatrix}\begin{Bmatrix} \mu\,\nu \\ \nu \end{Bmatrix}\right], \tag{12}$$

wo γ^{ik} die kontravarianten Komponenten des fünfdimensionalen metrischen Fundamentaltensors sind und wo $\begin{Bmatrix} r\,s \\ i \end{Bmatrix}$ die Christoffelschen Dreiindizessymbole bezeichnen, also:

$$\begin{Bmatrix} r\,s \\ i \end{Bmatrix} = \frac{1}{2}\sum \gamma^{i\mu}\left(\frac{\partial\gamma_{\mu r}}{\partial x^s} + \frac{\partial\gamma_{\mu s}}{\partial x^r} - \frac{\partial\gamma_{rs}}{\partial x^\mu}\right). \tag{13}$$

In dem Ausdruck von P denken wir uns, daß alle Größen von x^0 unabhängig sind, und daß $\gamma_{00} = \alpha$ ist.

Betrachten wir nun das über ein geschlossenes Gebiet des fünfdimensionalen Raumes ausgeführte Integral:

$$J = \int P \sqrt{-\gamma}\, d x^0\, d x^1\, d x^2\, d x^3\, d x^4, \tag{14}$$

wo γ die Determinante der γ_{ik} bedeutet.

Wir bilden δJ durch Variieren der Größen γ_{ik} und $\frac{\partial \gamma_{ik}}{\partial x^l}$, wobei deren Randwerte nicht verändert werden sollen. Hierbei soll α als eine Konstante betrachtet werden. Das Variationsprinzip:

$$\delta J = 0 \tag{15}$$

führt dann zu den folgenden Gleichungen:

$$R^{ik} - \frac{1}{2} g^{ik} R + \frac{\alpha \beta^2}{2} S^{ik} = 0 \quad (i, k = 1, 2, 3, 4) \tag{16a}$$

und

$$\frac{\partial \sqrt{-g}\, F^{i\mu}}{\partial x^\mu} = 0 \quad (i = 1, 2, 3, 4), \tag{16b}$$

wo R die Einsteinsche Krümmungsinvariante, R^{ik} die kontravarianten Komponenten des Einsteinschen Krümmungstensors, g^{ik} die kontravarianten Komponenten des Einsteinschen Fundamentaltensors, S^{ik} die kontravarianten Komponenten des elektromagnetischen Energie-Impulstensors, g die Determinante der g_{ik} und schließlich $F^{i\mu}$ die kontravarianten Komponenten des elektromagnetischen Feldtensors bedeuten. Setzen wir

$$\frac{\alpha \beta^2}{2} = \varkappa, \tag{17}$$

wo $\varkappa$ die von Einstein gebrauchte Gravitationskonstante bedeutet, so sehen wir, daß die Gleichungen (16a) in der Tat mit den Gleichungen der Relativitätstheorie für das Gravitationsfeld und (16b) mit den generalisierten Maxwellschen Gleichungen der Relativitätstheorie identisch sind für einen materiefreien Feldpunkt [1]).

Wenn wir uns auf die in der Elektronentheorie und der Relativitätstheorie übliche schematische Behandlungsweise der Materie beschränken, können wir die gewöhnlichen Gleichungen für den nicht materiefreien Fall in ähnlicher Weise erhalten. Wir ersetzen P in (14) durch

$$P + \varkappa \Sigma \gamma_{ik} \Theta^{ik}.$$

Um die Θ^{ik} zu definieren, wollen wir erst den auf ein Elektron oder einen Wasserstoffkern bezüglichen Tensor:

$$\vartheta^{ik} = \frac{d x^i}{d l} \frac{d x^k}{d l} \tag{18}$$

[1]) Siehe z. B. W. Pauli, Relativitätstheorie, S. 719 und 724.

betrachten, wo $d x^i$ die Lageänderungen des Teilchens bezeichnen, und $d l$ ein gewisses invariantes Differential bedeutet. Die Θ^{ik} sollen gleich der auf die Volumeneinheit bezogenen Summe der ϑ^{ik} für die verschiedenen Teilchen sein. Wir kommen dann wieder zu Gleichungen vom gewöhnlichen Typus, die mit den gewöhnlichen Feldgleichungen identisch werden, wenn wir setzen:

$$v_0 \frac{d\tau}{dl} = \pm \frac{e}{\beta c}, \tag{19}$$

$$\frac{d\tau}{dl} = \begin{cases} \sqrt{M} \\ \sqrt{m} \end{cases}, \tag{20}$$

wo allgemein

$$v_i = \Sigma \gamma_{i\mu} \frac{d x^\mu}{dl} \tag{21}$$

die kovarianten Komponenten des fünfdimensionalen Geschwindigkeitsvektors v^i sind, wo

$$v^i = \frac{d x^i}{dl}. \tag{22}$$

Ferner bedeuten e das elektrische Elementarquantum, M und m die Massen von Wasserstoffkern bzw. Elektron. Dabei gilt das obere Wertsystem für den Kern, das untere für das Elektron. Weiter ist

$$d\tau = \frac{1}{c} \sqrt{- d s^2}$$

das Differential der Eigenzeit.

Aus den Feldgleichungen folgen natürlich auf gewöhnliche Weise die Bewegungsgleichungen für materielle Teilchen und die Kontinuitätsgleichung. Die Rechnungen, die dazu führen, können von unserem Standpunkt aus einfach zusammengefaßt werden. Wie man leicht sieht, sind nämlich unsere Feldgleichungen mit den folgenden 14 Gleichungen äquivalent:

$$P^{ik} - \tfrac{1}{2}\gamma^{ik} P + \varkappa \Theta^{ik} = 0 \tag{23}$$

($i, k = 0, 1, 2, 3, 4$, aber nicht beide Null), wo die P^{ik} die kontravarianten Komponenten des verjüngten fünfdimensionalen Krümmungstensors sind (den R^{ik} entsprechend). Die in Frage stehenden Gleichungen folgen nun durch Divergenzbildung von (23). Hieraus folgt, daß sich die elektrischen Teilchen auf fünfdimensionalen geodätischen Linien bewegen, die den Bedingungen (19) und (20) genügen[1]). Wie man

[1]) Die speziellen Werte von $\frac{d\tau}{dl}$ sind natürlich in diesem Zusammenhang ohne Bedeutung. Wesentlich ist hier nur $\frac{d\tau}{dl} = \text{const.}$

sofort sieht, sind diese Bedingungen eben deshalb mit den Gleichungen der geodätischen Linien verträglich, weil x^0 in den γ_{ik} nicht vorkommt.

Es muß hier daran erinnert werden, daß wohl keine genügenden Gründe für die exakte Gültigkeit der Einsteinschen Feldgleichungen vorliegen. Immerhin möchte es nicht ohne Interesse sein, daß sich sämtliche 14 Feldgleichungen in so einfacher Weise vom Standpunkt der Theorie von Kaluza zusammenfassen lassen.

§ 2. Die Wellengleichung der Quantentheorie. Wir gehen nun dazu über, die Theorie der stationären Zustände und die damit zusammenhängenden charakteristischen Abweichungen von der Mechanik, die in der neueren Quantentheorie zum Vorschein kommen, in Beziehung zu der fünfdimensionalen Relativitätstheorie zu bringen. Betrachten wir zu diesem Zweck die folgende Differentialgleichung, die sich auf unseren fünfdimensionalen Raum beziehen soll und als eine einfache Verallgemeinerung der Wellengleichung betrachtet werden kann:

$$\sum a^{ik}\left(\frac{\partial^2 U}{\partial x^i \partial x^k} - \sum \begin{Bmatrix} i\,k \\ r \end{Bmatrix} \frac{\partial U}{\partial x^r}\right) = 0. \tag{24}$$

Hier bedeuten die a^{ik} die kontravarianten Komponenten eines fünfdimensionalen symmetrischen Tensors, die gewisse Funktionen der Koordinaten sein sollen. Die Gleichung (24) besteht unabhängig vom Koordinatensystem.

Betrachten wir erst eine durch (24) bestimmte Wellenausbreitung, die dem Grenzfall der geometrischen Optik entspricht. Wir kommen dazu, wenn wir setzen:

$$u = A\,e^{i\omega\Phi} \tag{25}$$

und ω als so groß annehmen, daß in (24) nur die mit ω^2 proportionalen Glieder berücksichtigt zu werden brauchen. Wir bekommen dann:

$$\sum a^{ik}\frac{\partial \Phi}{\partial x^i}\frac{\partial \Phi}{\partial x^k} = 0, \tag{26}$$

eine Gleichung, die der Hamilton-Jacobischen partiellen Differentialgleichung der Mechanik entspricht. Setzen wir

$$p_i = \frac{\partial \Phi}{\partial x^i}, \tag{27}$$

so können die Differentialgleichungen der Strahlen bekanntlich in der folgenden Hamiltonschen Form geschrieben werden:

$$\frac{d p_i}{-\dfrac{\partial H}{\partial x^i}} = \frac{d x^i}{\dfrac{\partial H}{\partial p_i}} = d\lambda, \tag{28}$$

wo
$$H = \frac{1}{2} \sum a^{ik} p_i p_k. \tag{29}$$

Aus (26) folgt noch
$$H = 0. \tag{30}$$

Eine andere Darstellung dieser Gleichungen, die der Lagrangeschen Form entspricht, ergibt sich durch den Umstand, daß die Strahlen als geodätische Nullinien der Differentialform:
$$\sum a_{ik} \, d x^i \, d x^k$$
betrachtet werden können, wo die a_{ik} die zu den a^{ik} reziproken Größen bedeuten, also
$$\sum a_{i\mu} a^{k\mu} = \delta_i^k = \begin{cases} 1, & i = k \\ 0, & i \neq k \end{cases}. \tag{31}$$

Setzen wir nun
$$\sum a_{ik} \, d x^i \, d x^k = \mu (d \vartheta)^2 + d s^2, \tag{32}$$
so können wir durch passende Wahl der Konstante μ erreichen, daß unsere Strahlengleichungen mit den Bewegungsgleichungen elektrischer Teilchen identisch werden. Setzen wir, um dies einzusehen:
$$L = \frac{1}{2} \mu \left(\frac{d \vartheta}{d \lambda}\right)^2 + \frac{1}{2} \left(\frac{d s}{d \lambda}\right)^2, \tag{33}$$
so folgt
$$p_0 = \frac{\partial L}{\partial \dfrac{d x^0}{d \lambda}} = \mu \frac{d \vartheta}{d \lambda} \tag{34}$$
und
$$p_i = \frac{\partial L}{\partial \dfrac{d x_i}{d \lambda}} = u_i \frac{d \tau}{d \lambda} + \beta p_0 \varphi_i \quad (i = 1, 2, 3, 4), \tag{35}$$
wo $u_1 \ldots u_4$ die kovarianten Komponenten des gewöhnlichen Geschwindigkeitsvektors bedeuten.

Die Strahlengleichungen lauten nun:
$$\frac{d p_0}{d \lambda} = 0, \tag{36 a}$$
$$\frac{d p_i}{d \lambda} = \frac{1}{2} \frac{\partial g_{\mu\nu}}{\partial x^i} \frac{d x^\mu}{d^\lambda} \frac{d x^\nu}{d \lambda} + \beta p_0 \frac{\partial \varphi_\mu}{\partial x^i} \frac{d x^\mu}{d \lambda} \quad (i = 1, 2, 3, 4). \tag{36}$$

Aus
$$\mu \, d \vartheta^2 + d s^2 = \mu \, d \vartheta^2 - c^2 d \tau^2 = 0$$
ergibt sich
$$\mu \frac{d \vartheta}{d \tau} = c \sqrt{\mu}. \tag{37}$$

Da nach (34) und (36 a) $\frac{d\vartheta}{d\lambda}$ und also auch $\frac{d\tau}{d\lambda}$ konstant ist, können wir λ so wählen, daß

$$\frac{d\tau}{d\lambda} = \begin{cases} M \text{ für den Wasserstoffkern,} \\ m \text{ für das Elektron.} \end{cases} \tag{38}$$

Ferner müssen wir, um zu den gewöhnlichen Bewegungsgleichungen zu gelangen, annehmen:

$$\beta p_0 = \begin{cases} +\frac{e}{c} \text{ für den Wasserstoffkern,} \\ -\frac{e}{c} \text{ für das Elektron.} \end{cases} \tag{39}$$

Aus (37) ergibt sich dann:

$$\mu = \begin{cases} \frac{e^2}{\beta^2 M^2 c^4} \text{ für den Wasserstoffkern,} \\ \frac{e^2}{\beta^2 m^2 c^4} \text{ für das Elektron.} \end{cases} \tag{40}$$

Die Gleichungen (35), (36) stimmen dann mit den gewöhnlichen Bewegungsgleichungen elektrischer Teilchen in Gravitationsfeldern und elektromagnetischen Feldern vollständig überein. Insbesondere sind die nach (35) definierten Größen p_i identisch mit den auf gewöhnliche Weise definierten generalisierten Momenten, was für die folgenden Überlegungen wichtig ist. Da wir β noch beliebig wählen können, wollen wir setzen:

$$\beta = \frac{e}{c}. \tag{41}$$

Es ergibt sich dann einfach:

$$p_0 = \begin{cases} +1 \text{ für den Wasserstoffkern,} \\ -1 \text{ für das Elektron,} \end{cases} \tag{39 a}$$

und

$$\mu = \begin{cases} \frac{1}{M^2 c^2} \text{ für den Wasserstoffkern,} \\ \frac{1}{m^2 c^2} \text{ für das Elektron.} \end{cases} \tag{40 a}$$

Wie man sieht, müssen wir in (37) für die Quadratwurzel das positive Zeichen im Falle des Kerns und das negative Zeichen im Falle des Elektrons wählen. Dies ist ja wenig befriedigend. Die Tatsache aber, daß man bei einem einzigen Werte von μ zwei verschiedene Klassen von Strahlen erhält, die sich gewissermaßen wie die positiven und negativen elektrischen Teilchen zueinander verhalten, könnte als ein Hinweis darauf angesehen werden, daß es vielleicht möglich ist, die Wellengleichung so abzuändern, daß sich die Bewegungsgleichungen beider Arten von Teilchen aus einem einzigen Wertsystem der Koeffizienten ergeben. Auf diese Frage wollen wir jetzt nicht weiter eingehen, sondern wir wollen dazu übergehen, die aus (32) folgende Wellengleichung im Falle des Elektrons etwas näher zu betrachten.

Da für das Elektron $p_0 = -1$ angenommen wurde, müssen wir nach (27) setzen:

$$\Phi = -x^0 + S(x^1, x^2, x^3, x^4). \tag{42}$$

Die Theorie von de Broglie ergibt sich nun, wenn wir die mit der Wellengleichung verträglichen, einem bestimmten Wert von ω entsprechenden stehenden Schwingungen aufsuchen, und dabei annehmen, daß die Wellenausbreitung nach den Gesetzen der geometrischen Optik vor sich geht. Dazu bedürfen wir des wohlbekannten Satzes von der Erhaltung der Phase, der sich sofort aus (28 und (30) ergibt. Es folgt nämlich:

$$\frac{d\Phi}{d\lambda} = \sum \frac{\partial \Phi}{\partial x^i}\frac{d x^i}{d\lambda} = \sum p_i \frac{\partial H}{\partial p_i} = 2H = 0. \tag{43}$$

Die Phase wird also von der Welle mitgeführt. Betrachten wir nun den einfachen Fall, wo sich Φ in zwei Teile spalten läßt, von denen der eine Teil nur von einer einzigen Koordinate, sagen wir x, abhängt, die mit der Zeit periodisch hin und her schwingt. Es wird dann eine stehende Schwingung möglich sein, die dadurch charakterisiert wird, daß eine in einem gewissen Augenblick durch (25) dargestellte harmonische Welle nach einer Periode von x mit derjenigen Welle in Phase zusammentrifft, die sich aus derselben Lösung (25) durch Einsetzen der neuen Werte von x^0, x^2, x^3, x^4 ergibt. Wegen der Erhaltung der Phase ist die Bedingung dafür einfach:

$$\omega \oint p\, d x = n \,.\, 2\pi, \tag{44}$$

wo n eine ganze Zahl bedeutet. Setzen wir:

$$\omega = \frac{2\pi}{h}, \tag{45}$$

wo h die Plancksche Konstante bedeutet, so ergibt sich also die gewöhnliche Quantenbedingung für eine separierbare Koordinate. Ähnliches gilt natürlich für ein beliebiges Periodizitätssystem. Die gewöhnliche Quantentheorie der Periodizitätssysteme entspricht also vollständig der Behandlung der Interferenzerscheinungen mittels der Annahme, daß sich die Wellen nach den Gesetzen der geometrischen Optik ausbreiten. Es mag noch hervorgehoben werden, daß wegen (42) die Beziehungen (44), (45) bei den Koordinatentransformationen (2) invariant bleiben.

Betrachten wir nun auch die Gleichung (24) in dem Falle, wo ω nicht so groß ist, daß wir nur die in ω quadratischen Glieder zu berücksichtigen brauchen. Wir beschränken uns dabei auf den einfachen Fall eines elektrostatischen Feldes. Dann haben wir in kartesischen Koordinaten:

$$\left.\begin{aligned} d\vartheta &= dx^0 - eV\,dt, \\ ds^2 &= dx^2 + dy^2 + dz^2 - c^2\,dt^2. \end{aligned}\right\} \tag{46}$$

Also ergibt sich:

$$H = \frac{1}{2}(p_x^2 + p_y^2 + p_z^2) - \frac{1}{2c^2}(p_t + eVp_0)^2 + \frac{m^2c^2}{2}p_0^2. \tag{47}$$

In der Gleichung (24) können wir nun die mit $\left\{ {ik \atop r} \right\}$ proportionalen Größen vernachlässigen, denn die Dreiindizessymbole sind in diesem Falle nach (17) kleine mit der Gravitationskonstante $\varkappa$ proportionale Größen. Wir bekommen also [1]):

$$\varDelta U - \frac{1}{c^2}\frac{\partial^2 U}{\partial t^2} - \frac{2eV}{c^2}\frac{\partial^2 U}{\partial t\,\partial x^0} + \left(m^2c^2 - \frac{e^2V^2}{c^2}\right)\frac{\partial^2 U}{\partial x^{0\,2}} = 0. \tag{48}$$

Für U können wir, da V nur von x, y, z abhängt, in Übereinstimmung mit (42) und (45) ansetzen:

$$U = e^{-2\pi i\left(\frac{x^0}{h} - \nu t\right)}\,\psi(x, y, z). \tag{49}$$

Dies in (48) eingeführt, ergibt:

$$\varDelta\psi + \frac{4\pi^2}{c^2h^2}\left[(h\nu - eV)^2 - m^2c^4\right]\psi = 0. \tag{50}$$

Setzen wir noch:

$$h\nu = mc^2 + E, \tag{51}$$

so bekommen wir die von Schrödinger [2]) gegebene Gleichung, deren stehende Schwingungen bekanntlich Werten von E entsprechen, die mit

---

[1]) Außer durch das Auftreten von x^0, das ja für die Anwendungen unwesentlich ist, unterscheidet sich diese Gleichung von der Schrödingerschen Gleichung durch die Art, in welcher in (48) die Zeit auftritt. Als eine Stütze für diese Form der Quantengleichung kann angeführt werden, daß dieselbe in dem Fall, wo V harmonisch von der Zeit abhängt, wie man durch eine einfache Störungsrechnung zeigen kann, Lösungen besitzt, die sich in ähnlicher Weise zu der Dispersionstheorie von Kramers verhalten wie die Schrödingerschen Lösungen zu der Quantentheorie der Spektrallinien. Diese Bemerkung verdanke ich Dr. W. Heisenberg.

[2]) Schrödinger, l. c.

den aus der Heisenbergschen Quantentheorie berechneten Energiewerten identisch sind. Wie man sieht, ist E in dem Grenzfall der geometrischen Optik gleich der auf gewöhnliche Weise definierten mechanischen Energie. Die Frequenzbedingung besagt, wie Schrödinger hervorgehoben hat, nach (51), daß die zu dem System gehörenden Lichtfrequenzen den aus den verschiedenen Werten der Frequenz ν gebildeten Differenzen gleich sind.

§ 3. Schlußbemerkungen. Wie die Arbeiten von de Broglie sind obenstehende Überlegungen aus dem Bestreben entstanden, die Analogie zwischen Mechanik und Optik, die in der Hamiltonschen Methode zum Vorschein kommt, für ein tieferes Verständnis der Quantenerscheinungen auszunutzen. Daß dieser Analogie ein reeller physikalischer Sinn zukommt, scheint ja die Ähnlichkeit der Bedingungen für die stationären Zustände von Atomsystemen mit den Interferenzerscheinungen der Optik anzudeuten. Nun stehen bekanntlich Begriffe wie Punktladung und materieller Punkt schon der klassischen Feldphysik fremd gegenüber. Auch wurde ja öfters die Hypothese ausgesprochen, daß die materiellen Teilchen als spezielle Lösungen der Feldgleichungen aufzufassen sind, welche das Gravitationsfeld und das elektromagnetische Feld bestimmen. Es liegt nahe, die genannte Analogie zu dieser Vorstellung in Beziehung zu bringen. Denn nach dieser Hypothese ist es ja nicht so befremdend, daß die Bewegung der materiellen Teilchen Ähnlichkeiten aufweist mit der Ausbreitung von Wellen. Die in Rede stehende Analogie ist jedoch unvollständig, solange man eine Wellenausbreitung in einem Raum von nur vier Dimensionen betrachtet. Dies kommt schon in der variablen Geschwindigkeit der materiellen Teilchen zum Vorschein. Denkt man sich aber die beobachtete Bewegung als eine Art Projektion auf den Zeitraum von einer Wellenausbreitung, die in einem Raum von fünf Dimensionen stattfindet, so läßt sich, wie wir sahen, die Analogie vollständig machen. Mathematisch ausgedrückt heißt dies, daß die Hamilton-Jacobische Gleichung nicht als Charakteristikengleichung einer vierdimensionalen, wohl aber einer fünfdimensionalen Wellengleichung aufgefaßt werden kann. In dieser Weise wird man zu der Theorie von Kaluza geführt.

Obwohl die Einführung einer fünften Dimension in unsere physikalischen Betrachtungen von vornherein befremdend sein mag, wird eine radikale Modifikation der den Feldgleichungen zugrunde gelegten Geometrie doch wieder in ganz anderer Weise durch die Quantentheorie nahegelegt. Denn es ist bekanntlich immer weniger wahrscheinlich geworden, daß die Quantenerscheinungen eine einheitliche raumzeitliche Beschreibung zulassen, wogegen die Möglichkeit, diese Erscheinungen durch ein System von fünfdimensionalen Feldgleichungen darzustellen, wohl nicht von vornherein auszuschließen ist [1]). Ob hinter diesen Andeutungen von Möglichkeiten etwas Wirkliches besteht, muß natürlich die Zukunft entscheiden. Jedenfalls muß betont werden, daß die in dieser Note versuchte Behandlungsweise, sowohl was die Feldgleichungen als auch die Theorie der stationären Zustände betrifft, als ganz provisorisch zu betrachten ist. Dies kommt wohl besonders in der auf S. 898 erwähnten schematischen Behandlungsweise der Materie zum Vorschein, sowie in dem Umstand, daß die zwei Arten von elektrischen Teilchen durch verschiedene Gleichungen vom Schrödingerschen Typus behandelt werden. Auch wird die Frage ganz offen gelassen, ob man sich bei der Beschreibung der physikalischen Vorgänge mit den 14 Potentialen begnügen kann, oder ob die Schrödingersche Methode die Einführung einer neuen Zustandsgröße bedeutet.

Mit den in dieser Note mitgeteilten Überlegungen habe ich mich sowohl in dem Physikalischen Institut der University of Michigan, Ann Arbor, wie in dem hiesigen Institut für theoretische Physik beschäftigt. Ich möchte auch an dieser Stelle Prof. H. M. Randall und Prof. N. Bohr meinen wärmsten Dank aussprechen.

---

1) Bemerkungen dieser Art, die Prof. Bohr bei mehreren Gelegenheiten gemacht hat, haben einen entschiedenen Einfluß auf das Entstehen der vorliegenden Note gehabt.

Chapter 9
Oskar Klein (1927): Zur fünfdimensionalen Darstellung der Relativitätstheorie

Oskar Klein. Zur fünfdimensionalen Darstellung der Relativitätstheorie. *Zeitschrift fur Physik*, 46: 188–208.

Zur fünfdimensionalen Darstellung der Relativitätstheorie.

Von O. Klein in Kopenhagen.

(Eingegangen am 22. Oktober 1927.)

Die sogenannte fünfte Dimension wird als der elektrischen Ladung konjugiert im Sinne der Quantentheorie aufgefaßt. Die Nichtbeobachtbarkeit der fünften Dimension wird hierdurch mit der Existenz des elektrischen Elementarquantums in Zusammenhang gebracht. Es wird versucht, eine einfache einheitliche Darstellung der Gesetze der Relativitätstheorie, insbesondere der fünf Erhaltungssätze für Impuls, Energie und Elektrizitätsmenge zu geben mit Hilfe der zuerst von Kaluza vorgeschlagenen fünfdimensionalen Riemannschen Geometrie.

Einleitung. Durch die Entwicklung der Quantentheorie haben bekanntlich die Begriffe Raum und Zeit den unmittelbaren Sinn verloren, der ihnen in der älteren Physik und auch in der Relativitätstheorie zugeschrieben wurde. Dies hängt, wie besonders Bohr* hervorgehoben hat, schon damit zusammen, daß unsere Meßwerkzeuge (Licht und freie materielle Teilchen) nach der Quantentheorie ein Verhalten zeigen, welches durch das Dilemma Korpuskel–Welle charakterisiert wird, wodurch eine scharfe Definition einer raumzeitlichen Koinzidenz, wie sie der gewöhnlichen Raum—Zeitauffassung zugrunde liegt, unmöglich wird. Von diesem Gesichtspunkt aus ist zu erwarten, daß die allgemeine Relativitätstheorie einer Revision im Sinne des Quantenpostulats bedürftig ist, wie dies auch daraus hervorgeht, daß verschiedene ihrer Konsequenzen mit den Forderungen der Quantentheorie im Widerspruch sind**.

Es zeigt sich nun bei der Betrachtung der Definitions- und Beobachtungsmöglichkeiten ein eigentümliches reziprokes Verhältnis der Raum–Zeitbeschreibung zu der Erhaltung von Energie und Impuls***. Dies hängt auf das innigste damit zusammen, daß nach de Broglie den freien materiellen Teilchen ebenso wie der Strahlung im leeren Raum — die einzigen Fälle, wo es bisher möglich war, in der Quantentheorie den

* N. Bohr, Die Naturwissenschaften. Im Erscheinen. Vgl. auch W. Heisenberg, ZS. f. Phys. **43**, 172, 1927.

** Zum Beispiel würde die Einsteinsche Strahlenablenkung an einer Anzahl im Zustand statistischen Temperaturgleichgewichts sich bewegender materieller Körper das Strahlungsgleichgewicht stören, was eine Art gravitationellen Comptoneffekt vermuten läßt. Ein entsprechendes Resultat wird man — schon durch das Nichtauftreten der Planckschen Konstante in den betreffenden Gleichungen — im Fall eines statistischen Gleichgewichts zwischen Gravitationswellen und Lichtwellen nach den Gesetzen der allgemeinen Relativitätstheorie erwarten.

*** N. Bohr, l. c.

Forderungen der Relativitätstheorie streng zu genügen — eine Wellenphase zuzuordnen ist, worin Raum—Zeitkoordinaten und Komponenten des Impuls—Energievektors in symmetrischer Weise auftreten. In dieser Weise erscheinen die Erhaltungssätze, im Anschluß an die Vorstellungen der Lichtquantentheorie, als verallgemeinerte Frequenzbedingung, wie dies besonders durch die Erfahrungen beim Comptoneffekt nahegelegt wird.

Wenn man an die scheinbar etwas zufällige Rolle der Erhaltungssätze in der analytischen Mechanik, als vier der vielen möglichen Integrale der Bewegungsgleichungen, denkt, so könnte es nach dem eben erwähnten Gesichtspunkt scheinen, als ob die Raum—Zeitgrößen, die für unsere unmittelbaren Erfahrungen von so zentraler Bedeutung sind, in dem mathematischen Gebäude der zukünftigen Quantentheorie in den Hintergrund treten würden. Ungeachtet der großen Fruchtbarkeit der Quantenmechanik spricht indessen vieles dafür, daß der Ausgangspunkt für eine allgemeine mathematische Formulierung der Quantentheorie eher in den Feldgleichungen der klassischen Theorie zu suchen ist, als in den Bewegungsgleichungen der Punktmechanik. Hierdurch würde die Frage nach der Rolle der Raum—Zeitgrößen (oder vielmehr ihres quantentheoretischen Analogons) in ein anderes Licht rücken, denn ein großer Teil der für die Feldtheorie charakteristischen Zustandsgrößen (die Einsteinschen Gravitationspotentiale) sind ja formal unmittelbar an die Raum—Zeitbegriffe geknüpft.

Im Hinblick auf diese Verhältnisse, also einerseits die Auflösung der gewöhnlichen Raum—Zeitauffassung durch die Quantentheorie und andererseits die anscheinende Notwendigkeit durch sinngemäße Umschreibung der gewöhnlichen Feldtheorie und im nächsten Anschluß an die Erhaltungssätze für Impuls und Energie die Begriffe Raum und Zeit neu zu begründen, liegt es nahe, den fünften allgemeinen Erhaltungssatz, die Erhaltung der Elektrizitätsmenge, als Ausgangspunkt für die Einführung einer neuen der elektrischen Ladung konjugierten Größe zu betrachten, die den vier raum—zeitlichen Koordinaten als formal gleichwertig an die Seite zu treten hat.

Die Existenz eines elektrischen Elementarquantums führt hierbei eine eigentümliche Situation mit sich, die beim ersten Anblick scheinbar gegen eine Zuordnung der Erhaltung der Elektrizität zu den übrigen Erhaltungssätzen spricht. Durch die naheliegende Annahme, daß die Existenz der Elementarladung die Folge der Quantisierung einer in dem entsprechenden klassischen Modell kontinuierlich veränderlichen Größe ist, verliert jedoch dieses Argument seine Beweiskraft. Würde doch auch, bei bestimmten, der Relativitätstheorie nicht widersprechenden Annahmen, die Impulskomponente in irgend einer Richtung nur diskrete Werte annehmen.

Die Tatsache, daß ein freies elektrisches Teilchen stets eine bestimmte Ladung hat, hat aber nach der hier vertretenen Auffassung die prinzipielle Nichtbeobachtbarkeit der neuen Dimension zur Folge. Dies ist eine einfache Folgerung aus der Dirac-Jordanschen statistischen Formulierung der Quantentheorie, die besonders klar in der von Bohr begründeten Auffassung der Zusammenhänge zwischen Definition und Beobachtung zum Ausdruck kommt. Diese Nichtbeobachtbarkeit der neuen Dimension, die mit der Tatsache, daß unsere Erfahrungen — abgesehen von der aus dem Quantenpostulat herrührenden Begrenzung — sich mit Hilfe der Begriffe Raum und Zeit ordnen lassen, übereinstimmt darf nicht als Beweis der Überflüssigkeit des Begriffs der fünften Dimension betrachtet werden. Als schlagendes Beispiel für die Nützlichkeit einer nicht-beobachtbaren Größe in der Quantenmechanik sei die Eigenrotation des Elektrons erwähnt, die klassisch mit Hilfe eines Winkels beschrieben wird, welcher in der Quantentheorie prinzipiell unbeobachtbar ist.

Wie zuerst Kaluza hervorgehoben hat, wird durch die Einführung einer fünften Dimension in die Relativitätstheorie in der Tat eine innige Verschmelzung der Gesetze des Elektromagnetismus mit denen der Gravitation erzielt, wodurch die Zustandsgrößen, welche das elektromagnetische Feld (die elektromagnetischen Potentiale) charakterisieren, in einer ähnlichen Beziehung zu den metrischen Größen erscheinen wie die Einsteinschen Gravitationspotentiale in der gewöhnlichen Form der Relativitätstheorie. Hierdurch und nach dem erwähnten Gesichtspunkt scheint sich diese fünfdimensionale Form der Relativitätstheorie als der natürliche Ausgangspunkt für eine allgemeine Quantenfeldtheorie darzubieten. Im Hinblick hierauf soll im folgenden ein Versuch zu einer zusammenfassenden Darstellung der Grundlagen dieser Form der Relativitätstheorie versucht werden, wobei der Rahmen der klassischen Theorie — abgesehen von einigen einfachen Betrachtungen zur Wellenmechanik — noch nicht verlassen wird. Es handelt sich also um eine nur in formaler Hinsicht von der gewöhnlichen Relativitätstheorie sich unterscheidende Darstellung*.

§ 1. Der Erhaltungssatz bei den elektrischen Teilchen. An den in der Einleitung entwickelten Gesichtspunkt anknüpfend, wollen wir

* Das folgende bildet eine sehr verspätete ausführlichere Darstellung der Bestrebungen des Verfassers zur fünfdimensionalen Form der Relativitätstheorie,

zuerst die Erhaltungssätze in dem einfachen Fall eines elektrisch geladenen Teilchens betrachten. Die Erhaltung von Energie und Impuls wird hier durch die Bewegungsgleichungen ausgedrückt, während die Erhaltung der Elektrizität einfach aussagt, daß die Ladung des Teilchens konstant bleibt. Nehmen wir zuerst an, daß kein elektromagnetisches Feld vorhanden ist. Die Bewegung geht dann nach Einstein auf geodätischen Linien der Raum—Zeitmannigfaltigkeit vor sich. Es sei die Lage des Teilchens durch die vier Raum—Zeitkoordinaten x^1, x^2, x^3, x^4 definiert. Das zu dem Teilchen gehörige Differential der Eigenzeit ist dann gegeben durch

$$d\tau = \sqrt{-\frac{g_{ik}}{c^2}\,d x^i\,d x^k},$$

wo c die Vakuumlichtgeschwindigkeit bedeutet, und die g_{ik} die zehn Komponenten des metrischen Fundamentaltensors bezeichnen, die nach Einstein sowohl die metrischen als auch die gravitationellen Eigenschaften des Raumes zum Ausdruck bringen. Führen wir die kovarianten Komponenten des Impulsenergievektors des Teilchens p_1, p_2, p_3, p_4 ein, wo

$$p_i = m\,g_{ik}\,\frac{d x^k}{d\tau} \qquad (i = 1, 2, 3, 4) \tag{1}$$

und m die Masse des Teilchens bedeutet, so lassen sich die Bewegungsgleichungen, d. h. die Gleichungen der geodätischen Linien, bekanntlich auf die folgende Hamiltonsche Form bringen:

$$\frac{d x^i}{d\tau} = \frac{\partial H_0}{\partial p_i}, \quad \frac{d p_i}{d\tau} = -\frac{\partial H_0}{\partial x^i}, \quad (i = 1, 2, 3, 4), \tag{2}$$

wo

$$H_0 = \frac{1}{2m}\,g^{ik}\,p_i\,p_k. \tag{3}$$

---

die — da sich inzwischen verschiedene Forscher mit diesem Thema beschäftigt haben — von mathematischen Resultaten wohl kaum neues bringt. Ich hoffe, daß dieselbe doch etwas Interesse beanspruchen darf, da sie sich sowohl durch die physikalischen Gesichtspunkte wie durch die Form von den übrigen Arbeiten — auch den früheren Arbeiten des Verfassers — unterscheidet. Insbesondere halte ich es nicht mehr für möglich, durch die Einführung einer fünften Dimension den von der Quantentheorie geforderten Abweichungen von der Raum—Zeitbeschreibung der klassischen Theorie gerecht zu werden. Hier sei auf die von de Broglie gegebene Übersicht über den Gegenstand hingewiesen (Journ. de phys. **8**, 65, 1927). Ausführliche Literaturangaben findet man bei L. Rosenfeld, Acad. Roy. de Belgique (5) **13**, Heft 6, 1927. Vgl. außerdem H. Mandel (ZS. f. Phys. **39**, 136, 1926; **45**, 285, 1927), der unabhängig von Kaluza und dem Verfasser zu der fünfdimensionalen Darstellung gekommen ist, sowie F. Gonseth und G. Juvet, C. R. **185**, 341, 412, 448, 535, 1927.

Wir fügen nun zu den Koordinaten x^1, x^2, x^3, x^4 eine weitere fünfte Koordinate x^0 hinzu und zu den Impulsenergiekomponenten p_1, p_2, p_3, p_4 eine zu x^0 konjugierte Größe p_0, die gemäß dem in der Einleitung Gesagten proportional der Ladung e des Teilchens angenommen werden soll. Setzen wir

$$p_0 = \frac{e}{\beta c}, \tag{4}$$

wo β eine später zu bestimmende universelle Konstante bezeichnet von der Dimension einer elektrischen Ladung dividiert durch eine Energie. Zu den Gleichungen (2) können wir also das Paar hinzufügen

$$\frac{d x^0}{d \tau} = \frac{\partial H_0}{\partial p_0} = 0, \quad \frac{d p_0}{d \tau} = -\frac{\partial H_0}{\partial p_0} = 0 \tag{5}$$

als Ausdruck der Konstanz der elektrischen Ladung.

Wir fragen demnächst nach der Form der entsprechenden Gleichungen bei der Anwesenheit eines elektromagnetischen Feldes, das wir durch die vier Komponenten φ_1, φ_2, φ_3, φ_4 des elektromagnetischen Potentialvektors charakterisieren wollen. Bekanntlich lassen sich auch in diesem Falle die Bewegungsgleichungen auf Hamiltonsche Form bringen, wenn wir für die Hamiltonsche Funktion setzen

$$H = \frac{1}{2m} g^{ik} \left(p - \frac{e}{c} \varphi_i\right) \left(p_k - \frac{e}{c} \varphi_k\right) + \text{const.} \tag{6}$$

In diesem Ausdruck ersetzen wir e/c durch βp_0 und nehmen aus später ersichtlichen Gründen die beliebige additive Konstante in H gleich $\frac{1}{2m} p_0^2$ an. Den resultierenden Ausdruck können wir folgendermaßen schreiben

$$H = \frac{1}{2m} \{g^{ik} p_i p_k - 2 \beta \varphi^k p_k p_0 + (1 + \beta^2 \varphi_k \varphi^k) p_0^2\}, \tag{7}$$

wo die φ^k die kontravarianten Komponenten des Potentialvektors bedeuten. Wie wir sehen, hat dieser Ausdruck eine entsprechende Form wie der Ausdruck (3) für die zu den geodätischen Linien gehörige Hamiltonsche Funktion. Dies führt uns dazu, versuchsweise eine solche Metrik des fünfdimensionalen Raumes einzuführen, daß die kontra varianten Komponenten γ^{ik} des Fundamentaltensors gegeben sind durch

$$\gamma^{ik} = g^{ik}, \quad \gamma^{i0} = -\beta \varphi^i, \quad \gamma^{00} = 1 + \beta^2 \varphi_k \varphi^k \quad (i, k = 1, 2, 3, 4). \tag{8}$$

Wir können nämlich dann schreiben

$$H = \frac{1}{2m} \sum \gamma^{ik} p_i p_k, \tag{9}$$

wo das Summenzeichen wie überall im folgenden eine Summation von 0 bis 4 über doppelt auftretende Indizes andeuten soll. Die kanonischen Gleichungen

$$\frac{d x^i}{d\tau} = \frac{\partial H}{\partial p_i}, \quad \frac{d p_i}{d\tau} = -\frac{\partial H}{\partial x^i} \quad (i = 0, 1, 2, 3, 4) \tag{10}$$

sind nun die Gleichungen fünfdimensionaler geodätischer Linien, von denen das erste Paar die Erhaltung der Ladung (wegen des Nichtvorkommens von x^0 in H) bzw. die Definition von $\frac{d x^0}{d\tau}$ ausdrücken, während die vier übrigen Paare die Bewegungsgleichungen darstellen.

Wenn γ_{ik} die kovarianten Komponenten des Fundamentaltensors bezeichnen, so ergibt sich aus (9) auf Grund von (10)

$$p_i = m \sum \gamma_{ik} \frac{d x^k}{d\tau}. \tag{11}$$

Wenn andererseits die Form (7) von H benutzt wird, so folgt aus (10)

$$m \frac{d x^0}{d\tau} = (1 + \beta^2 \varphi_k \varphi^k) p_0 - \beta \varphi^k p_k, \quad m \frac{d x^i}{d\tau} = g^{ik} p_k - \beta \varphi^i p_0$$
$$(i = 1, 2, 3, 4)$$

oder

$$p_i = m g_{ik} \frac{d x^k}{d\tau} + \beta p_0 \varphi_i \quad (i = 1, 2, 3, 4) \tag{12}$$

und also auch

$$p_0 = m \left\{ \frac{d x^0}{d\tau} + \beta \varphi_k \frac{d x^k}{d\tau} \right\},$$

$$p_i = m \left\{ \beta \varphi_i \frac{d x^0}{d\tau} + (g_{ik} + \beta^2 \varphi^i \varphi_k) \frac{d x^k}{d\tau} \right\} (i, k = 1, 2, 3, 4), \tag{13}$$

woraus durch Vergleich mit (11) folgt

$$\gamma_{ik} = g_{ik} + \beta^2 \varphi_i \varphi_k, \quad \gamma_{i0} = \beta \varphi_i, \quad \gamma_{00} = 1 \quad (i, k = 1, 2, 3, 4). \tag{14}$$

Aus diesen Werten für die γ_{ik} ersieht man, daß dem Quadrat $d\sigma^2$ des fünfdimensionalen Linienelements die folgende Form gegeben werden kann

$$d\sigma^2 = d\vartheta^2 + d s^2, \tag{15}$$

wo

$$d s^2 = g_{ik} d x^i d x^k \tag{16}$$

das Quadrat des **Einstein**schen Linienelements bezeichnet, während

$$d\vartheta = d x^0 + \beta \varphi_i d x^i. \tag{17}$$

Aus der letzten Form des Linienelements können wir leicht das gewonnene Resultat verifizieren, daß die Bewegungsgleichungen zusammen mit der Erhaltung der Ladung durch die Gleichungen fünfdimensionaler geodätischer Linien dargestellt werden können. Es lassen sich nämlich die zu (15) gehörigen geodätischen Linien durch die folgenden Lagrangeschen Gleichungen darstellen:

$$\frac{d}{d\tau}\left(\frac{\partial L}{\partial \frac{d x^i}{d\tau}}\right) = \frac{\partial L}{\partial x^i} \quad (i = 0,\, 1,\, 2,\, 3,\, 4), \tag{18}$$

wo

$$L = \frac{1}{2}\, m \left(\frac{d\sigma}{d\tau}\right)^2 = \frac{m}{2}\left\{g_{ik}\frac{d x^i}{d\tau}\,\frac{d x^k}{d\tau} + \left(\frac{d x^0}{d\tau} + \beta\,\varphi_i\,\frac{d x^i}{d\tau}\right)^2\right\}. \tag{19}$$

Betrachten wir nun außerdem die folgende Größe

$$L_0 = \frac{m}{2} g_{ik}\frac{d x^i}{d\tau}\,\frac{d x^k}{d\tau} + \beta\, p_0\, \varphi_i \frac{d x^i}{d\tau}, \tag{20}$$

so sehen wir leicht, daß wenn x irgend eine der Größen x^1, x^2, x^3, x^4, $\frac{d x^1}{d\tau}$, $\frac{d x^2}{d\tau}$, $\frac{d x^3}{d\tau}$, $\frac{d x^4}{d\tau}$ bedeutet, $\frac{\partial L}{\partial x} = \frac{\partial L_0}{\partial x}$ ist. In den vier der Gleichungen (18), welche die Bewegung des Teilchens ausdrücken, dürfen wir also L durch L_0 ersetzen, bekommen aber dann, wegen $\beta\, p_0 = e/c$, die bekannte Lagrangesche Form der Bewegungsgleichungen eines elektrischen Teilchens.

Für elektrische Teilchen in einem gravitationellen und elektromagnetischen Felde lassen sich also die fünf Erhaltungssätze in die Aussage zusammenfassen: Die „Bahn“ eines elektrischen Teilchens in der fünfdimensionalen Mannigfaltigkeit ist eine geodätische Linie des Linienelements (15). Die Konstanz der elektrischen Ladung erscheint hierbei als Folge des Nichtvorkommens von x^0 in den Komponenten des fünfdimensionalen metrischen Fundamentaltensors.

Es mag bemerkt werden, daß die in der gewöhnlichen Form der Relativitätstheorie fremdartig anmutende Form der Momente

$$p_i = m\, u_i + \frac{e}{c}\,\varphi_i,$$

wo u_1, u_2, u_3, u_4 die kovarianten Komponenten der Vierergeschwindigkeit bedeuten [vgl. Gleichung (12)], in der fünfdimensionalen Darstellung in einem natürlichen Lichte erscheint. Nach (11) hat man in der Tat

$$p_i = m \sum \gamma_{ik} \frac{d x^k}{d\tau},$$

wo

$$\sum \gamma_{ik} \frac{d x^k}{d\tau}$$

offenbar eine in bezug auf die fünfdimensionale Metrik kovariante Komponente der „Fünfergeschwindigkeit“ bedeutet.

§ 2. Invarianzeigenschaften der betrachteten fünfdimensionalen Mannigfaltigkeit. Nach den Ergebnissen des vorigen Paragraphen werden wir versuchen, die Gesetzmäßigkeiten der Relativitätstheorie allgemein mit Hilfe einer fünfdimensionalen Riemannschen Geometrie darzustellen. Da wir fürs erste nicht den Rahmen der gewöhnlichen Relativitätstheorie verlassen wollen, ist es zweckmäßig, hierbei die Koordinatenwahl so zu beschränken, daß vier der fünf Koordinaten (sagen wir x^1, x^2, x^3, x^4) sich stets auf die in wohlbekannter Weise definierte Raum–Zeitmannigfaltigkeit der Relativitätstheorie beziehen. Um zu den gewöhnlichen physikalischen Gesetzen zu gelangen, müssen wir die fundamentale Annahme über das Linienelement $d\sigma$ des fünfdimensionalen Raumes machen, daß bei geeigneter Wahl der fünften Koordinate die Größen γ_{ik} nur von den Raum–Zeitkoordinaten x^1, x^2, x^3, x^4 abhängen. Wir werden dann die Koordinatenwahl in bezug auf x^0 so einschränken, daß diese Bedingung in allen zum Gebrauch gelangenden Koordinatensystemen erfüllt ist. Dies mag als ein — wohl vorläufiger — Ausdruck für die in der Einleitung erwähnte Nichtbeobachtbarkeit der fünften Dimension betrachtet werden.

Betrachten wir nun eine durch die folgenden Gleichungen ausgedrückte Koordinatentransformation

$$x^{0'} = \psi(x^0, x^1, x^2, x^3, x^4), \quad x^{i'} = \psi_i(x^1, x^2, x^3, x^4) \qquad (i = 1, 2, 3, 4),$$

so folgt

$$d x^{0'} = \sum \frac{\partial \psi}{\partial x^i}\, d x^i = \sum \alpha^i\, d x^i.$$

Die Größen $\alpha_i = \dfrac{\partial \psi}{\partial x^i}$ müssen aber auf Grund des eben Gesagten von x^0 unabhängig sein, da diese Koordinate sonst in den γ'_{ik} auftreten würde. Man hat nun

$$\frac{\partial \alpha_0}{\partial x^i} = \frac{\partial \alpha_i}{\partial x^0} \quad (i = 1, 2, 3, 4).$$

Folglich muß α_0 eine Konstante sein, die wir stets gleich Eins annehmen wollen, da diese weitere Einschränkung, wie wir gleich sehen werden, die Gleichheit einer gegebenen elektrischen Ladung in allen erlaubten Bezugssystemen zur Folge hat. Durch unsere Annahmen reduzieren sich also die zugelassenen Koordinatentransformationen auf die folgende Gruppe

$$x^{0'} = x^0 + \psi_0(x^1, x^2, x^3, x^4), \quad x^{i'} = \psi_i(x^1, x^2, x^3, x^4) \quad (i = 1, 2, 3, 4). \qquad \text{(A)}$$

Wir wollen die Invarianzeigenschaften der in Betracht kommenden Vektoren und Tensoren der Gruppe (A) gegenüber untersuchen. Es seien $\sigma^0, \sigma^1, \sigma^2, \sigma^3, \sigma^4$ die kontravarianten Komponenten eines fünfdimensionalen Vektors. Wir bilden die Komponenten desselben Vektors in dem gestrichenen Koordinatensystem. Für $i = 1, 2, 3, 4$ hat man

$$\sigma^{i'} = \sum \frac{\partial x^{i'}}{\partial x^{\mu}} \sigma^{\mu} = \frac{\partial x^{i'}}{\partial x^{\mu}} \sigma^{\mu}.$$

Die Komponenten $\sigma^1, \sigma^2, \sigma^3, \sigma^4$ verhalten sich folglich wie die kontravarianten Komponenten eines Vierervektors. Betrachten wir auch die kovariante Komponente σ_0. Es ist

$$\sigma_0 = \sum \frac{\partial x^{\mu'}}{\partial x^0} \sigma'_{\mu} = \sigma'_0.$$

Die Größe σ_0 ist also eine Invariante, woraus nach (4) die Invarianz der Ladung folgt.

Betrachten wir demnächst einen fünfdimensionalen Tensor zweiten Ranges mit den Komponenten Θ^{ik}. Wir haben dann für $i, k = 1, 2, 3, 4$

$$\Theta^{i k'} = \sum \frac{\partial x^{i'}}{\partial x^{\mu}} \frac{\partial x^{k'}}{\partial x^{\nu}} \Theta^{\mu\nu} = \frac{\partial x^{i'}}{\partial x^{\mu}} \frac{\partial x^{k'}}{\partial x^{\nu}} \Theta^{\mu\nu}.$$

Die raum-zeitlichen kontravarianten Komponenten des Tensors bilden demnach einen vierdimensionalen Tensor. Für die gemischten Komponenten $\Theta_0^{\,i}$, wo $i = 1, 2, 3, 4$, hat man ferner

$$\Theta_0^{\,i'} = \sum \Theta_{\mu}^{\,\nu} \frac{\partial x^{i'}}{\partial x^{\nu}} \frac{\partial x^{\mu}}{\partial x^{0'}} = \Theta_0^{\,\nu} \frac{\partial x^{i'}}{\partial x^{\mu}}.$$

Diese bilden also einen Vierervektor, und das gleiche gilt von den Komponenten $\Theta^i_{\ 0}$. Die Komponente Θ_{00} bildet eine Invariante, denn es gilt

$$\Theta_{00} = \sum \frac{\partial x^{\mu'}}{\partial x^0} \frac{\partial x^{\nu'}}{\partial x^0} \Theta'_{\mu\nu} = \Theta'_{00}.$$

Schließlich betrachten wir die Größe $\sum \Theta_{0i}\, d x^i$. Da

$$\Theta_{0i} = \sum \frac{\partial x^{\mu'}}{\partial x^0} \frac{\partial x^{\nu'}}{\partial x^i} \Theta'_{\mu\nu} = \sum \frac{\partial x^{\nu'}}{\partial x^i} \Theta'_{0\nu},$$

so hat man

$$\sum \Theta_{0i}\, d x^i = \sum \Theta'_{0\nu} \frac{\partial x^{\nu'}}{\partial x^i} d x^i = \sum \Theta'_{0\nu}\, d x^{\nu'}.$$

Auch diese Größe bildet folglich eine Invariante, was natürlich auch von der Größe $\sum \Theta_{i0}\, d x^i$ gilt.

Wenden wir diese Resultate auf den Fundamentaltensor γ_{ik} an, so folgt erstens, daß die Festsetzung $\gamma^{ik} = g^{ik}$ für $i, k = 1, 2, 3, 4$ mit den Transformationen (A) verträglich ist. Aus dem Umstand, daß $\sum \gamma_{0i}\, d x^i$ eine Invariante bildet, folgt ferner, wegen der Invarianz von γ_{00}, daß die vier Größen γ_{0i} sich invariantentheoretisch wie die elektromagnetischen Potentiale verhalten. In der Tat sind diese nur so weit bestimmt, daß die Größe $\varphi_i\, d x^i$ bis auf ein vollständiges Differential einer beliebigen Funktion von x^1, x^2, x^3, x^4 festgelegt ist. Sowohl die Festsetzungen $\gamma_{i0} = \beta\,\varphi_i$ wie $\gamma_{00} = 1$ sind also mit unserer Transformationsgruppe verträglich*. Da wir identisch schreiben können

$$d\sigma^2 = \sum \gamma_{ik}\, d x^i\, d x^k \equiv \frac{1}{\gamma_{00}} (\sum \gamma_{0k}\, d x^k)^2 + \left(\gamma_{ik} - \frac{\gamma_{i0}\,\gamma_{k0}}{\gamma_{00}}\right) d x^i\, d x^k,$$

so sehen wir auch auf Grund der Invarianz von γ_{00} und $\sum \gamma_{0k}\, d x^k$, daß $\left(\gamma_{ik} - \frac{\gamma_{i0}\,\gamma_{k0}}{\gamma_{00}}\right) d x^i\, d x^k$ eine Invariante bildet, was wiederum seine in (14) ausgedrückte Identifizierung mit dem Quadrat des Einsteinschen Linienelements rechtfertigt.

Betrachten wir noch einmal den Vektor $\sigma^0, \sigma^1, \sigma^2, \sigma^3, \sigma^4$. Setzen wir

$$s^i = \sigma^i \quad (i = 1, 2, 3, 4), \tag{21}$$

und drücken wir die kovarianten Komponenten s_1, s_2, s_3, s_4 des Vierervektors mit Hilfe der kovarianten Komponenten $\sigma_0, \sigma_1, \sigma_2, \sigma_3, \sigma_4$ des Fünfvektors aus. Man hat

$$s_i = g_{ik}\, s^k = g_{ik} \sum \gamma^{kr}\, \sigma_r = \sigma_i - \beta\,\varphi_i\,\sigma^0, \tag{22}$$

eine Formel, von der (12) ein Spezialfall ist. Setzen wir ferner bei einem symmetrischen Tensor

$$\Theta^{ik} = T^{ik} \quad (i, k = 1, 2, 3, 4),$$

und berechnen wir ebenso die Komponenten T_{ik}. Es gilt

$$\begin{aligned} T_{ik} &= g_{i\mu}\, g_{k\nu}\, \Theta^{\mu\nu} = g_{i\mu}\, g_{k\nu} \sum \gamma^{\mu r} \gamma^{\nu s}\, \Theta_{rs} \\ &= \Theta_{ik} - \beta\,(\varphi_i\,\Theta_{0k} + \varphi_k\,\Theta_{0i}) + \beta^2\,\varphi_i\,\varphi_k\,\Theta_{00}. \end{aligned} \tag{23}$$

Von dieser Formel ist wieder der aus (14) folgende Ausdruck für die g_{ik} ein Spezialfall.

---

* Wie man sieht, verlangt die im vorigen Paragraphen besprochene einfache Deutung der Bewegungsgleichungen, daß γ_{00} konstant ist. In § 4 werden wir einem Argument für das positive Zeichen von γ_{00} begegnen, d. h. für die Raumartigkeit der fünften Dimension. Die Annahme $\gamma_{00} = 1$ bedeutet dann nur, daß wir den Maßstab für x^0 in entsprechender Weise wie für die übrigen Dimensionen wählen. Auf die Frage der Konstanz von γ_{00} werden wir noch in § 4 zurückkehren.

Schließlich wollen wir die Determinante $\gamma = |\gamma_{ik}|$ der Größen γ_{ik} berechnen. Wir haben

$$\gamma^{-1} = \begin{vmatrix} 1+\beta^2\varphi_k\varphi^k, & -\beta\varphi^1, & -\beta\varphi^2, & -\beta\varphi^3, & -\beta\varphi^4 \\ -\beta\varphi^1 & g^{11} & g^{12} & g^{13} & g^{14} \\ -\beta\varphi^2 & g^{21} & g^{22} & g^{23} & g^{44} \\ -\beta\varphi^3 & g^{31} & g^{32} & g^{33} & g^{34} \\ -\beta\varphi^4 & g^{41} & g^{42} & g^{43} & g^{44} \end{vmatrix} = (1+\beta^2\varphi_k\varphi^k)\,g^{-1} - \beta^2\varphi^i\varphi^k g_{ik} g^{-1} = g^{-1},$$

wo g die Determinante der g_{ik} bezeichnet. Wir bekommen also einfach

$$\gamma = g. \tag{24}$$

§ 3. Das allgemeine Erhaltungsprinzip. Wir fangen damit an, einen einfachen mathematischen Hilfssatz abzuleiten. Durch eine infinitesimale Koordinatentransformation mögen die Koordinaten x^0, x^1, x^2, x^3, x^4 in $\bar{x}^0, \bar{x}^1, \bar{x}^2, \bar{x}^3, \bar{x}^4$ verwandelt werden, wobei

$$\bar{x}^i = x^i + \varepsilon\,\xi^i(x^1, x^2, x^3, x^4) \qquad (i = 0, 1, 2, 3, 4). \tag{25}$$

Hier soll ε eine unendlich kleine Konstante bedeuten, während die ξ^i Funktionen der vier Koordinaten x^1, x^2, x^3, x^4 seien, die am Rande eines gewissen geschlossenen Gebietes G des fünfdimensionalen Raumes verschwinden sollen. Wir vergleichen die Werte der Größen g_{ik}, φ_i vor und nach der Transformation in zwei Punkten, wo die alten und neuen Koordinaten die gleichen Werte haben. Eine solche Variation bezeichnen wir mit δ^* †. Bedeutet a irgend eine der zu variierenden Größen, so hat man also

$$\delta^* a = \bar{a}(x^0, x^1, x^2, x^3, x^4) - a(x^0, x^1, x^2, x^3, x^4), \tag{26}$$

wo $\bar{a}$ die transformierte Größe a bedeutet.

Es seien nun Θ^{ik} die Komponenten eines symmetrischen fünfdimensionalen Tensors, die x^0 nicht enthalten sollen. Dann zeigt man leicht, daß

$$\sum \Theta^{ik}\delta\gamma_{ik} = \Theta^{ik}\delta g_{ik} + 2\beta\,\Theta_0^k\,\delta\varphi_k, \tag{27}$$

wo δ irgend eine beliebige infinitesimale Variation der durch (14) gegebenen Größen γ_{ik} bedeutet. Sei nun δ speziell die Variation δ^*, so folgt nach (25) und (26) durch eine leichte Rechnung

$$\tfrac{1}{2}\int\cdots\int\sqrt{-\gamma}\,dx^0\ldots dx^4\sum\Theta^{ik}\delta^*\gamma_{ik} = \varepsilon\int\cdots\int\sqrt{-\gamma}\,dx^0\ldots dx^4\sum \varDelta\mathrm{iv}_i\Theta\,.\,\xi^i, \tag{28}$$

wo die Integration über das erwähnte Gebiet G zu erstrecken ist, und wo

$$\varDelta\mathrm{iv}_i\Theta \equiv \frac{1}{\sqrt{-\gamma}}\sum\frac{\partial\sqrt{-\gamma}\,\Theta_i^k}{\partial x^k} - \frac{1}{2}\sum\frac{\partial\gamma_{rs}}{\partial x^i}\,\Theta^{rs}. \tag{29}$$

† Vgl. W. Pauli, Relativitätstheorie (Leipzig-Berlin 1921), S. 617.

Ebenso folgt, wenn wir den vierdimensionalen Tensor mit den Komponenten $T^{ik} = \Theta^{ik}$ $(i, k = 1, 2, 3, 4)$ mit T bezeichnen,

$$\frac{1}{2}\int\cdots\int\sqrt{-g}\,d\,x^0\cdots d\,x^4\,T^{ik}\,\delta^* g_{ik}$$
$$= \varepsilon\int\cdots\int\sqrt{-g}\,d\,x^0\cdots d\,x^4\,\mathrm{Div}_i T\,.\,\xi^i, \tag{30}$$

wo

$$\mathrm{Div}_i T \equiv \frac{1}{\sqrt{-g}}\frac{\partial\sqrt{-g}\,T_i^k}{\partial x^k} - \frac{1}{2}\frac{\partial g_{rs}}{\partial x^i}T^{rs}. \tag{31}$$

Ferner findet man leicht

$$\int\cdots\int\sqrt{-\gamma}\,d\,x^0\cdots d\,x^4\sum\Theta_0^i\,\delta^*\gamma_{0i}$$
$$= \varepsilon\beta\int\cdots\int\sqrt{-\gamma}\,d\,x^0\cdots d\,x^4\left\{\left(F_{ik}\,\Theta_0^i + \frac{1}{\sqrt{-\gamma}}\frac{\partial\sqrt{-\gamma}\,\Theta_0^i}{\partial x^i}\varphi_k\right)\xi^k\right.$$
$$\left. + \frac{1}{\beta\sqrt{-\gamma}}\frac{\partial\sqrt{-\gamma}\,\Theta_0^i}{\partial x^i}\xi^0\right\}, \tag{32}$$

wo

$$F_{ik} = \frac{\partial\varphi_k}{\partial x^i} - \frac{\partial\varphi_i}{\partial x^k} \tag{33}$$

die Komponenten des elektromagnetischen Feldtensors bezeichnen. Da nun die ξ^i innerhalb des Gebietes G ganz beliebig gewählt werden können, folgt aus (22), da $\gamma = g$, identisch

$$\Delta\mathrm{iv}_0\,\Theta \equiv \frac{1}{\sqrt{-g}}\frac{\partial\sqrt{-g}\,\Theta_0^k}{\partial x^k},\quad \Delta\mathrm{iv}_i\,\Theta \equiv \mathrm{Div}_i T - \beta F_{ik}\,\Theta_0^k$$
$$+ \frac{\beta}{\sqrt{-g}}\frac{\partial\sqrt{-g}\,\Theta_0^k}{\partial x^k}\varphi_i \quad (i = 1, 2, 3, 4). \tag{34}$$

Wir gehen nun dazu über, die Erhaltungssätze zu betrachten, wie sie in der Relativitätstheorie dargestellt werden. Es seien s^1, s^2, s^3, s^4 die Komponenten des Viererstroms; dann kann der Satz von der Erhaltung der Elektrizität bekanntlich durch folgende Gleichung ausgedrückt werden

$$\frac{\partial\sqrt{-g}\,s^k}{\partial x^k} = 0. \tag{35}$$

Wenn ferner T^{ik} $(i, k = 1, 2, 3, 4)$ die Komponenten des Energie-Impulstensors der Materie vorstellen (das elektromagnetische Feld nicht mit einbegriffen), so kann das Energie-Impulsprinzip bekanntlich folgendermaßen ausgedrückt werden

$$\mathrm{Div}_i T = F_{ik}\,s^k \quad (i = 1, 2, 3, 4). \tag{36}$$

Nach dem vorigen Paragraphen können wir aber einen fünfdimensionalen symmetrischen Tensor durch folgende Gleichungen definieren

$$\Theta^{ik} = T^{ik},\ \Theta_0^k = \frac{1}{\beta} s^k \quad (i, k = 1, 2, 3, 4), \tag{37}$$

wobei es auf den Wert von Θ_{00} nicht ankommt. Es folgt dann aus (34)

$$\Delta\mathrm{iv}_0 \Theta = \frac{1}{\beta\sqrt{-g}} \frac{\partial \sqrt{-g}\, s^k}{\partial x^k},\ \Delta\mathrm{iv}_i \Theta = \mathrm{Div}_i T - F_{ik} s^k$$
$$+ \frac{1}{\sqrt{-g}} \frac{\partial \sqrt{-g}\, s^k}{\partial x^k} \varphi_i \quad (i = 1, 2, 3, 4). \tag{38}$$

Die fünf Erhaltungssätze können folglich in die fünfdimensionale Divergenzbeziehung

$$\Delta\mathrm{iv}_i \Theta = 0 \quad (i = 0, 1, 2, 3, 4) \tag{39}$$

zusammengefaßt werden, deren erste Gleichung die Erhaltung der Elektrizität ausspricht, während die vier übrigen das Energie-Impulsprinzip ausdrücken. Diese Gleichung bildet das fünfdimensionale Analogon zu dem Energie-Impulsprinzip der gewöhnlichen Relativitätstheorie bei Abwesenheit elektromagnetischer Felder.

Als Erläuterung dieser Beziehungen betrachten wir zuerst den Beitrag zum Energie-Impulstensor, den nach der gewöhnlichen Relativitätsmechanik ein Einzelteilchen von der Masse m liefert. Dieser ist $m u^i u^k$, wo u^i die Komponenten der Vierergeschwindigkeit darstellen. Hat das Teilchen eine Ladung e, so liefert es weiter einen Beitrag $\frac{e}{c} u^i$ zu dem Viererstrom. Einem solchen Teilchen können wir nun nach (37) einen fünfdimensionalen Tensor, wir wollen ihn mit ϑ^{ik} bezeichnen, zuordnen, indem wir setzen

$$\vartheta^{ik} = m u^i u^k,\ \vartheta_0^k = \frac{e}{\beta c} u^k. \tag{40}$$

Bezeichnen wir nun die Fünfergeschwindigkeit mit v^0, v^1, v^2, v^3, v^4, wo (vgl. S. 194)

$$v^i = u^i \ (i = 1, 2, 3, 4),\ v_0 = \frac{p_0}{m} = \frac{e}{m\beta c}, \tag{41}$$

so können wir anstatt (40) schreiben

$$\vartheta^{ik} = m v^i v^k,\ \vartheta_0^k = m v_0 v^k, \tag{42}$$

woraus die fünfdimensionale Tensorform unmittelbar hervorgeht.

Leiten wir demnächst das sogenannte wellenmechanische Energie-Impulsprinzip ab. Zu diesem Zwecke betrachten wir die folgende Invariante

$$M = \frac{h^2}{8\pi^2 m}\sum \gamma^{ik}\frac{\partial U}{\partial x^i}\frac{\partial U}{\partial x^k} + \frac{1}{2}\lambda U^2, \tag{43}$$

wo h die Plancksche Konstante, m wieder die Masse des Teilchens, von dessen Wellengleichung die Rede ist, bedeuten, während

$$\lambda = m c^2 - \frac{e^2}{\beta^2 m c^2} \tag{44}$$

und U eine invariante Funktion von x^0, x^1, x^2, x^3, x^4 sein soll. Wir bilden die Variation $\delta(\sqrt{-\gamma}\,M)$, indem wir die Größen γ^{ik} und U variieren, und integrieren diese Größe über ein bestimmtes fünfdimensionales Gebiet, an dessen Rand die Variationen verschwinden mögen. Es folgt nach einer einfachen Rechnung

$$\int\cdots\int\delta(\sqrt{-\gamma}\,M)\,dx^0\cdots dx^4 = \int\cdots\int\sqrt{-\gamma}\,dx^0\cdots dx^4\Big\{\sum\Theta_{ik}\delta\gamma^{ik} + \Big(-\frac{h^2}{4\pi^2 m}\,\bigcirc\, U + \lambda U\Big)\delta U\Big\}, \tag{45}$$

wo

$$\Theta_{ik} = \frac{h^2}{8\pi^2 m}\frac{\partial U}{\partial x^i}\frac{\partial U}{\partial x^k} - \frac{1}{2}M\gamma_{ik} \quad (i, k = 0, 1, 2, 3, 4), \tag{46}$$

während $\bigcirc$ die fünfdimensionale Verallgemeinerung des Laplaceschen Operators bedeutet, also

$$\bigcirc\, U = \frac{1}{\sqrt{-\gamma}}\sum\frac{\partial}{\partial x^i}\Big(\sqrt{-\gamma}\,\gamma^{ik}\frac{\partial U}{\partial x^k}\Big). \tag{47}$$

Lassen wir nun speziell unsere Variation die durch (26) ausgedrückte infinitesimale Änderung bedeuten. Dann haben wir nach (28)

$$\tfrac{1}{2}\int\cdots\int\sqrt{-\gamma}\,dx^0\ldots dx^4\,\Sigma\,\Theta_{ik}\delta^*\gamma^{ik} = -\varepsilon\int\cdots\int\sqrt{-\gamma}\,dx^0\ldots dx^4\sum\Delta\mathrm{iv}_i\Theta\,.\,\xi^i \tag{28a}$$

und ferner, da

$$\overline{U}(\bar{x}^0, \bar{x}^1, \bar{x}^2, \bar{x}^3, \bar{x}^4) = U(x^0, x^1, x^2, x^3, x^4),$$

$$\delta^* U = -\varepsilon\sum\frac{\partial U}{\partial x^i}\xi^i, \tag{48}$$

so daß die rechte Seite von (45) die Form annimmt

$$-2\varepsilon\int\cdots\int\sqrt{-\gamma}\,dx^0\ldots dx^4\sum\xi^i\Big\{\Delta\mathrm{iv}_i\,\Theta + \frac{1}{2}\Big(-\frac{h^2}{4\pi^2 m}\,\bigcirc\, U + \lambda U\Big)\frac{\partial U}{\partial x^i}\Big\}.$$

Da aber $\int \ldots \int \sqrt{-\gamma}\, M\, d x^0 \ldots d x^4$ eine Invariante ist, so muß die Größe auf der linken Seite von (45) für beliebige ξ^i, die am Rande des Gebiets verschwinden, gleich Null sein*. Es folgen also die Identitäten

$$\varDelta \mathrm{iv}_i \Theta + \frac{1}{2}\left(-\frac{h^2}{4\pi^2 m} \pentagon U + \lambda U\right)\frac{\partial U}{\partial x^i} = 0. \tag{49}$$

In der fünfdimensionalen Schreibweise lautet aber die Wellengleichung für das elektrische Teilchen**

$$-\frac{h^2}{4\pi^2 m} \pentagon U + \lambda U = 0, \tag{50}$$

so daß folgt

$$\varDelta \mathrm{iv}_i \Theta = 0. \tag{51}$$

Diese Bezeichnungen drücken, wie wir zeigen wollen, das allgemeine wellenmechanische Erhaltungsprinzip aus. Wie aus (49) hervorgeht, ist dieses Prinzip nicht nur eine Folge der Wellengleichung (50), sondern die Wellengleichung folgt auch umgekehrt aus (51), so daß beide völlig äquivalent sind.

Wir setzen wieder $\Theta^{ik} = T^{ik}$ $(i, k = 1, 2, 3, 4)$ und berechnen mit Hilfe von (23) die Komponenten T_{ik}. Es ergibt sich nach einer leichten Rechnung

$$T_{ik} = \frac{h^2}{8\pi^2 m}\left(\frac{\partial U}{\partial x^i} - \beta \varphi_i \frac{\partial U}{\partial x^0}\right)\left(\frac{\partial U}{\partial x^k} - \beta \varphi_k \frac{\partial U}{\partial x^0}\right) - \frac{1}{2} M g_{ik} (i, k = 1, 2, 3, 4). \tag{52}$$

Setzen wir ferner $\Theta_0^k = \frac{1}{\beta} s^k$, so folgt nach (22)

$$s_i = \frac{h^2}{8\pi^2 m} \frac{\partial U}{\partial x^0}\left(\frac{\partial U}{\partial x^i} - \beta \varphi_i \frac{\partial U}{\partial x^0}\right) \quad (i = 1, 2, 3, 4). \tag{53}$$

Um zu der vierdimensionalen Form der Wellengleichung zu gelangen, müssen wir nun setzen

$$U = \varphi e^{\frac{2\pi i}{h} p_0 x^0} + \psi e^{-\frac{2\pi i}{h} p_0 x^0}, \tag{54}$$

* Vgl. W. Pauli, l. c.

** Diese Form der Wellengleichung hat zuerst de Broglie, l. c. vorgeschlagen. Wegen (54) ist sie der früher vom Verfasser (ZS. f. Phys. **37**, 895, 1926) und von Fock (ebenda **39**, 226, 1926) gegebenen Form äquivalent. Die von de Broglie geäußerte entgegengesetzte Ansicht beruht auf einer Verwechslung der auf S. 901 der Arbeit des Verfassers definierten Größen α_{ik} mit den metrischen Größen γ_{ik}. Doch sei betont, daß sich nur die Form (50) der Wellengleichung direkt an die in § 1 der vorliegenden Arbeit und S. 899 der früheren Arbeit gegebene Darstellung der Bewegungsgleichungen anschließt. Die Tatsache, daß sich der Grenzübergang zur klassischen Theorie mittels der älteren Form der Gleichung besonders einfach gestaltet, bietet — im Gegensatz zu meiner früheren Ansicht — jedoch kein Argument für diese Form der Gleichung dar, da die bezügliche Grenzbetrachtung wegen der Unteilbarkeit der elektrischen Ladung kaum als sinngemäß anzusehen ist.

wo φ und ψ zwei konjugiert komplexe nur von x^1, x^2, x^3, x^4 abhängige Funktionen bezeichnen. Durch Einführung dieses Ausdrucks in (52) und (53) zerfallen die Größen T_{ik} und s_i in drei Teile, einen von x^0 unabhängigen und zwei mit $e^{2\frac{2\pi i}{h}p_0 x^0}$ bzw. $e^{-2\frac{2\pi i}{h}p_0 x^0}$ proportionale Teile, die offensichtlich jeder für sich die Erhaltungsgleichungen (51) befriedigen müssen. Für den Gebrauch der Größen T_{ik} und S_i in der Quantentheorie kommen jedoch nur die von x^0 unabhängigen Teile in Betracht, die wir als Mittelwerte über x^0 mit $\overline{T}_{ik}$ bzw. $\overline{s}_i$ bezeichnen wollen. Es ergibt sich wegen $p_0 = \frac{e}{\beta c}$

$$\overline{T}_{ik} = \frac{1}{2m}\left\{\left(\frac{h}{2\pi i}\frac{\partial\varphi}{\partial x^i} - \frac{e}{c}\varphi_i\varphi\right)\left(-\frac{h}{2\pi i}\frac{\partial\psi}{\partial x^k} - \frac{e}{c}\varphi_k\psi\right)\right.$$
$$\left. + \left(\frac{h}{2\pi i}\frac{\partial\varphi}{\partial x^k} - \frac{e}{c}\varphi_k\varphi\right)\left(-\frac{h}{2\pi i}\frac{\partial\psi}{\partial x^i} - \frac{e}{c}\varphi_i\psi\right) - \frac{1}{2}\overline{M}g_{ik},\right\} \quad (55)$$

wo

$$\overline{M} = \frac{1}{m}g^{ik}\left(\frac{h}{2\pi i}\frac{\partial\varphi}{\partial x^i} - \frac{e}{c}\varphi_i\varphi\right)\left(-\frac{h}{2\pi i}\frac{\partial\psi}{\partial x^k} - \frac{e}{c}\varphi_k\psi\right) + mc^2\varphi\psi. \quad (56)$$

Weiter folgt

$$s_i = \frac{e}{2mc}\left\{\frac{h}{2\pi i}\left(\psi\frac{\partial\varphi}{\partial x^i} - \varphi\frac{\partial\psi}{\partial x^i}\right) - 2\frac{e}{c}\varphi_i\varphi\psi\right\} \quad (57)$$

in Übereinstimmung mit zuerst von Schrödinger und Gordon gegebenen Ausdrücken für den Energie-Impulstensor und den Viererstrom. Auf Grund von (51) und (34) müssen die Ausdrücke (55) und (57) tatsächlich den Erhaltungssätzen genügen.

§ 4. Die Feldgleichungen. Wir wollen jetzt von unserem Gesichtspunkt aus die Feldgleichungen der Relativitätstheorie betrachten. Es seien R_{ik} die Komponenten des verjüngten Krümmungstensors, also

$$R_{ik} \equiv \frac{\partial\Gamma^\alpha_{i\alpha}}{\partial x^k} - \frac{\partial\Gamma^\alpha_{ik}}{\partial x^\alpha} + \Gamma^\beta_{i\alpha}\Gamma^\alpha_{k\beta} - \Gamma^\alpha_{ik}\Gamma^\beta_{\alpha\beta} \quad (i, k = 1, 2, 3, 4), \quad (58)$$

wo

$$\Gamma^i_{rs} \equiv \frac{1}{2}g^{ik}\left(\frac{\partial g_{kr}}{\partial x^s} + \frac{\partial g_{ks}}{\partial x^r} - \frac{\partial g_{rs}}{\partial x^k}\right) \quad (i, r, s = 1, 2, 3, 4) \quad (59)$$

die wohlbekannten Dreiindizessymbole bezeichnen. Es sei ferner

$$G_{ik} \equiv R_{ik} - \frac{1}{2}g_{ik}R, \quad (60)$$

wo

$$R \equiv g^{ik}R_{ik} \quad (61)$$

die Krümmungsinvariante bedeutet. Mit diesen Bezeichnungen lauten bekanntlich die Einsteinschen Feldgleichungen für das Gravitationsfeld

$$G^{ik} = -\varkappa\,(T^{ik} + S^{ik}) \qquad (i, k = 1, 2, 3, 4), \tag{62}$$

wo wir aus später ersichtlichen Gründen die kontravarianten Komponenten gewählt haben. Hier sind S^{ik} die kontravarianten Komponenten des Energie-Impulstensors des elektromagnetischen Feldes, dessen kovariante Komponenten lauten

$$S_{ik} \equiv F_{ri}\,F_{sk}\,g^{rs} - \frac{1}{4}\,F_{rs}\,F^{rs}\,g_{ik} \qquad (i, k = 1, 2, 3, 4), \tag{63}$$

während $\varkappa$ die Einsteinsche Gravitationskonstante bedeutet.

Ferner lautet die Einsteinsche Verallgemeinerung der Maxwellschen Gleichungen für das elektromagnetische Feld

$$\frac{1}{\sqrt{-g}}\,\frac{\partial \sqrt{-g}\,F^{ik}}{\partial x^k} = s^i \qquad (i = 1, 2, 3, 4). \tag{64}$$

Bekanntlich gilt nun identisch

$$\mathrm{Div}_i\, G \equiv 0, \quad \frac{\partial \sqrt{-g}\,s^k}{\partial x^k} \equiv 0 \qquad (i = 1, 2, 3, 4), \tag{65}$$

wo wir einen Augenblick s^i als Abkürzung des Ausdrucks auf der linken Seite von (64) gebraucht haben. Ferner hat man

$$\mathrm{Div}_i\, S \equiv -F_{ik}\,s^k \qquad (i, k = 1, 2, 3, 4), \tag{66}$$

wo wiederum s^k als Abkürzung für $\frac{1}{\sqrt{-g}}\,\frac{\partial \sqrt{-g}\,F^{ik}}{\partial x^k}$ steht. Diese Identitäten führen auf Grund der Feldgleichungen (62) und (64) in wohlbekannter Weise zu den Erhaltungssätzen für elektrische Ladung, Energie und Impuls.

Nach diesen Vorbereitungen betrachten wir den fünfdimensionalen Tensor Γ^{ik}, der das genaue Analogon zu dem vierdimensionalen Tensor G^{ik} ist. Er hängt nach (14) von den g_{ik}, den φ_i und deren ersten und zweiten Differentialquotienten nach x^1, x^2, x^3, x^4 ab. Ferner erfüllt er die Identitäten

$$\Delta\mathrm{iv}_i\, \Gamma \equiv 0 \qquad (i = 0, 1, 2, 3, 4), \tag{67}$$

wofür wir nach (34) auch schreiben können

$$\frac{1}{\sqrt{-g}}\,\frac{\partial \sqrt{-g}\,\Gamma_0^k}{\partial x^k} \equiv 0, \quad \mathrm{Div}_i\, \overline{G} - \beta\,F_{ik}\,\Gamma_0^k \equiv 0 \qquad (i = 1, 2, 3, 4), \tag{68}$$

wo $\overline{G}$ einen vierdimensionalen Tensor mit den Komponenten Γ^{ik} ($i, k = 1, 2, 3, 4$) bezeichnet. Nach § 2 müssen ja diese Γ^{ik} der Gruppe (A) gegenüber einen symmetrischen vierdimensionalen Tensor bilden.

Betrachten wir zuerst die Komponenten Γ^k_0, die nach § 2 der Gruppe (A) gegenüber einen Vierervektor bilden, dessen Divergenz nach (68) identisch verschwindet. Aus den Größen g_{ik}, φ_i selbst und ihren ersten Differentialquotienten läßt sich, wie man leicht sieht, keine Größe bilden, die (A) gegenüber sich wie ein Vierervektor verhält. (Die φ_i bilden ja nur dann einen Vierervektor, wenn x^0 untransformiert bleibt.) Die einzigen unabhängigen Tensoren zweiten Ranges, die aus diesen Größen gebildet werden können, sind ferner g_{ik} und F_{ik}. Wenn wir noch die zweiten Differentialquotienten der g_{ik} und φ_i hinzunehmen, und zwar linear, so bekommen wir als einzigen Vierervektor const $\frac{1}{\sqrt{-g}} \frac{\partial \sqrt{-g}\, F^{ik}}{\partial x^k}$, dessen Divergenz wirklich identisch verschwindet. Von dieser Form muß also Γ^i_0 sein. Den konstanten Faktor können wir bestimmen, indem wir z. B. in dem allgemeinen Ausdruck für Γ^i_0 alle Größen g_{ik} und drei der Größen φ_i als konstant annehmen. Es folgt dann

$$\Gamma^i_0 = -\frac{1}{2}\beta \frac{1}{\sqrt{-g}} \frac{\partial \sqrt{-g}\, F^{ik}}{\partial x^k} \qquad (i = 1, 2, 3, 4). \tag{69}$$

Aus (68) und (66) folgt nun

$$\mathrm{Div}_i\, G \equiv \frac{1}{2}\beta^2\, \mathrm{Div}_i\, S \qquad (i = 1, 2, 3, 4). \tag{70}$$

Es muß also $\overline{G}_{ik} - \frac{1}{2}\beta^2 S_{ik}$ ein Tensor mit identisch verschwindender Divergenz sein. Die einzigen unabhängigen aus den g_{ik}, φ_i und deren ersten und zweiten Differentialquotienten zu bildenden Größen, die sich (A) gegenüber wie symmetrische Tensoren zweiten Ranges verhalten und außerdem linear sind in den zweiten Differentialquotienten, sind aber neben den schon genannten (g_{ik}, F_{ik}) nur G_{ik}, $R g_{ik}$, S_{ik} und $F_{rs} F^{rs} g_{ik}$. Verlangen wir noch, daß die Divergenz des Tensors identisch verschwinden soll, so bleibt als einzige Möglichkeit übrig const G_{ik} + const g_{ik}. Da nun durch Nullsetzen der φ_i, wie leicht zu zeigen ist, die Γ^{ik} einfach in die G^{ik} übergehen, so muß also gelten

$$\Gamma^{ik} = G^{ik} + \frac{1}{2}\beta^2 S^{ik} \qquad (i, k = 1, 2, 3, 4). \tag{71}$$

Wir wollen nun die Annahme machen, daß die noch unbestimmte Konstante β folgendermaßen mit der Einsteinschen Gravitationskonstante $\varkappa$ zusammenhängt*

$$\beta = \sqrt{2\varkappa}. \tag{72}$$

Wir können dann zufolge (37) den Feldgleichungen der Relativitätstheorie die folgende einfache Form geben

$$\Gamma^{ik} = -\varkappa\Theta^{ik},\ \Gamma_0^k = -\varkappa\Theta_0^k \quad (i, k = 1, 2, 3, 4), \tag{73}$$

wo Θ den fünfdimensionalen Impulsenergie-Elektrizitätstensor bedeutet.

Da es nicht notwendig ist, die Größe Θ_{00} festzulegen, können wir diesen Gleichungen eine noch symmetrischere Form geben, wenn wir als Definition von Θ_{00} noch die folgende Gleichung hinzufügen

$$\Gamma_{00} = -\varkappa\Theta_{00}. \tag{74}$$

Durch eine einfache Rechnung folgt hieraus

$$\Theta_{00} = \frac{1}{2} g_{ik}\Theta^{ik} + \frac{3}{4} F_{ik}F^{ik}, \tag{75}$$

und wir können für die Feldgleichungen schreiben

$$\Gamma_{ik} = -\varkappa\Theta_{ik} \quad (i, k = 0, 1, 2, 3, 4), \tag{76}$$

welches außer (75) eben 14 Bestimmungsgleichungen für die g_{ik} und φ_i ergibt.

Eine andere Weise, dem Fehlen einer fünfzehnten Feldgleichung Rechnung zu tragen, ergibt sich, wenn wir die Variation des Integrals $\int\ldots\int\sqrt{-\gamma}\,P\,dx^0\ldots dx^4$ betrachten, das über ein geschlossenes fünfdimensionales Gebiet G zu erstrecken ist, und wo P die fünfdimensionale Krümmungsinvariante bedeutet. Man hat nach bekannten Formeln, wenn die γ_{ik} variiert und am Rande des Gebiets festgehalten werden,

$$\delta\int\ldots\int\sqrt{-\gamma}\,P\,dx^0\ldots dx^4 = -\int\ldots\int\sqrt{-\gamma}\,dx^0\ldots dx^4\sum\Gamma^{ik}\delta\gamma_{ik}. \tag{77}$$

Geben wir nun von vornherein den γ_{ik} die Werte (14), so daß γ_{00} unvariiert bleibt, so folgt auf Grund von (27)

$$\delta\int\ldots\int\sqrt{-\gamma}\,P\,dx^0\ldots dx^4$$
$$= -\int\ldots\int\sqrt{-\gamma}\,dx^0\ldots dx^4\,\{\Gamma^{ik}\delta g_{ik} + 2\beta\,\Gamma_0^k\,\delta\varphi_k\}. \tag{78}$$

In derselben Weise folgt

$$\int\ldots\int\sqrt{-\gamma}\,dx^4\ldots dx^4\sum\Theta^{ik}\delta\gamma_{ik}$$
$$= -\int\ldots\int\sqrt{-\gamma}\,dx^0\ldots dx^4\,\{\Theta^{ik}\delta g_{ik} + 2\beta\,\Theta_0^k\,\delta\varphi_k\}. \tag{79}$$

* Hätten wir die fünfte Dimension als zeitartig betrachtet, so hätte das Glied $\frac{1}{2}\beta^2 S^{ik}$ das umgekehrte Zeichen bekommen, was als Argument für die Raumartigkeit der fünften Dimension zu betrachten ist.

Wenn C den nicht zu variierenden Wert von $\sum \gamma_{ik}\Theta^{ik}$ bedeutet, so wird der Ausdruck linker Hand in der letzten Gleichung offenbar gleich*

$$\delta\int\ldots\int\sqrt{-\gamma}\,dx^0\ldots dx^4\left\{\sum\gamma_{ik}\Theta^{ik}-C\right\},$$

wo die γ_{ik} zu variieren sind, so daß wir die Feldgleichungen (73) zusammenfassen können durch das Variationsprinzip

$$\delta\int\ldots\int\sqrt{-\gamma}\,dx^0\ldots dx^4\left\{P-\varkappa\left(\sum\gamma_{ik}\Theta^{ik}-C\right)\right\}\cdot \tag{80}$$

In ähnlicher Weise können die aus den wellenmechanischen Ansätzen (46) für den Tensor Θ^{ik} fließenden Feldgleichungen nach (45) durch das folgende Variationsprinzip dargestellt werden

$$\delta\int\ldots\int\sqrt{-\gamma}\,dx^0\ldots dx^4\,(P+\varkappa M)=0, \tag{81}$$

wo die Variation nach U noch zu der Wellengleichung (50) führt.

Ein anderer naheliegender Ausweg, um die volle fünfdimensionale Symmetrie der Feldgleichungen herbeizuführen, ist, die Annahme $\gamma_{00}=1$ fallen zu lassen. In der Tat möchte man versucht sein die fünfzehnte Größe γ_{00} in Verbindung zu bringen mit der die Materie charakterisierenden Wellenfunktion U, um so eine formale Vereinigung von Materie und Feld zu erreichen. Da ein Versuch in dieser Richtung keine sehr versprechenden Resultate ergeben hat, wollen wir uns hier damit begnügen die Ergebnisse der Rechnung kurz mitzuteilen.

Anstatt der vorhin benutzten γ_{ik} führen wir andere Größen $\bar\gamma_{ik}$ ein, die folgendermaßen mit den alten Größen zusammenhängen

$$\bar\gamma_{ik}=e^{\varLambda}\gamma_{ik}\qquad(i,k=0,1,2,3,4), \tag{82}$$

wo $\varLambda$ irgend eine Funktion von x^0, x^1, x^2, x^3, x^4 bedeutet. Die zu den $\bar\gamma_{ik}$ gehörige Krümmungsinvariante sei mit $\bar P$ bezeichnet. Dann ergibt sich

$$\bar P=e^{-\varLambda}\left\{P+4\,\bigcirc\,\varLambda+3\sum\gamma^{ik}\frac{\partial\varLambda}{\partial x^i}\frac{\partial\varLambda}{\partial x^k}\right\}. \tag{83}$$

Die zugehörigen Größen $\bar\Gamma^{ik}$ können wir aus der Variationsformel

$$\delta\int\ldots\int\sqrt{-\bar\gamma}\,dx^0\ldots dx^4\,\bar P=-\int\ldots\int\sqrt{-\bar\gamma}\,dx^0\ldots dx^4\sum\bar\Gamma^{ik}\delta\bar\gamma_{ik}$$

ableiten. Eine einfache Rechnung zeigt nun, daß

$$\int\ldots\int\sqrt{-\bar\gamma}\,dx^0\ldots dx^4\,\bar P$$
$$=\int\ldots\int\sqrt{-\gamma}\,dx^0\ldots dx^4\left\{\Phi^2P-\frac{16}{3}\sum\gamma^{\mu\nu}\frac{\partial\Phi}{\partial x^\mu}\frac{\partial\Phi}{\partial x^\nu}\right\}, \tag{84}$$

wo

$$\Phi=e^{3/4\,\varLambda}=(\bar\gamma_{00})^{3/4}. \tag{85}$$

* In dem früher vom Verfasser, l. c., S. 898 gegebenen Ausdruck fehlt die Größe C.

Wenn wir diese Formel mit (81) vergleichen, so scheint im ersten Augenblick mit Rücksicht auf (43) eine Ähnlichkeit zu bestehen. Eine nähere Betrachtung zeigt jedoch, daß diese Ähnlichkeit illusorisch ist, was besonders durch den Zeichenunterschied zum Vorschein kommt.

Solange wir uns überhaupt an die klassischen Theorien halten, dürfte die obige Darstellung mit der Annahme $\gamma_{00} = 1$, die in sich geschlossen ist und zu den Gesetzen der gewöhnlichen Relativitätstheorie führt, die befriedigenste sein. Eine allgemeinere Verwertung der fünfdimensionalen Feldgleichungen kann man wohl erst durch die Einführung der Quantentheorie erwarten, wobei auch die Annahme, daß die γ_{ik} von x^0 unabhängig sind, durch eine rationellere zu ersetzen sein wird*.

---

* Die Existenz des elektrischen Elementarquantums legt in der Tat die Annahme nahe, daß die γ_{ik} — wenn wir von der Einführung der Quantentheorie in die raum—zeitlichen Relationen vorläufig absehen — als Matrizen von der Form

$$\begin{pmatrix} \alpha_{ik}, & \beta_{ik}\, e^{\frac{2\pi i}{h} p_0 x^0} \\ \beta^*_{ik}\, e^{-\frac{2\pi i}{h} p_0 x^0} & \alpha^*_{ik} \end{pmatrix}$$

zu betrachten sind, wo α_{ik} und α^*_{ik}, β_{ik} und β^*_{ik} konjugiert komplex sind, eine Form, die sowohl bei Additionen wie auch bei Multiplikationen nach den Matrizenrechenregeln erhalten bleibt. Es liegt hier ein Fall vor, der eine gewisse Ähnlichkeit aufzuweisen scheint mit der von Pauli und Jordan gegebenen Behandlung des rotierenden Elektrons. London (Naturwiss. **15**, 5, 1927) hat auf die Möglichkeit hingewiesen, x^0 als Drehwinkel des Elektrons um seine Achse zu betrachten. Doch dürfte dies kaum zutreffend sein, da nach der Definition von x^0 die zu ihr konjugierte Größe die elektrische Ladung und nicht der Drehimpuls des Elektrons ist. Es scheint jedoch ein innerer Zusammenhang zu bestehen zwischen der eben vorgeschlagenen Einführung der fünften Dimension und der von Jordan in Analogie zu der Theorie des rotierenden Elektrons gegebenen Darstellung der antisymmetrischen Lösung der Koordinatenraumgleichung, worauf ich demnächst zurückzukommen hoffe.

Chapter 10
Hugo Tetrode (1928): Allgemein-relativistische Quantentheorie des Elektrons

Hugo Tetrode (1928). Allgemein-relativistische Quantentheorie des Elektrons. *Zeitschrift für Physik*, 50: 336–346.

Allgemein-relativistische Quantentheorie des Elektrons.

Von **H. Tetrode** in Amsterdam.

(Eingegangen am 19. Juni 1928.)

§ 1. Algebraische Beziehungen für die verallgemeinerten Diracschen Matrizen γ_μ. — § 2. Beschränkung der zuzulassenden kanonischen Transformationen. — § 3. Das Wirkungsprinzip und die für die γ_μ vorzuschreibende partielle Differentialgleichung. — § 4. Das Erhaltungsgesetz der Elektrizität und der Impuls-Energiesatz.

Im folgenden wird eine Erweiterung der Diracschen Theorie* vorgenommen, die der Forderung der allgemeinen Kovarianz genügt, und zugleich eine direktere Ableitung des Impuls-Energiesatzes gestattet, als auf dem Boden der speziellen Relativität möglich war**.

§ 1. Die Diracschen Bedingungsgleichungen für die vier vierreihigen Matrizen γ_μ:

$$\gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu = 2\,\delta_\mu^\nu \qquad (\delta_\mu^\nu = 1 \text{ für } \mu = \nu \text{ und } = 0 \text{ für } \mu \neq \nu), \tag{1}$$

ersetzen wir durch

$$\gamma_\mu \gamma_\nu + \gamma_\nu \gamma_\mu = 2\,g_{\mu\nu}, \tag{2}$$

wo $g_{\mu\nu}$ die Komponenten des metrischen Fundamentaltensors sind. Wie in der klassischen Theorie sollen sie stets „gewöhnliche" Größen sein, d. h. sie sollen sich von der Einheitsmatrix nur durch einen Faktor unterscheiden und sind daher mit allen Matrizen aus der Gruppe der γ_μ bei der Multiplikation vertauschbar. Nach (2) sind die γ_μ ebenso wie die $g_{\mu\nu}$ im allgemeinen Raumzeitfunktionen, während sie früher als Konstanten betrachtet werden konnten***. Neben den kovarianten Vektorkomponenten γ_μ definieren wir kontravariante γ^μ durch

$$\gamma^\mu = \sum_\nu g^{\mu\nu} \gamma_\nu, \tag{3}$$

wo die $g^{\mu\nu}$ die aus der allgemeinen Relativitätstheorie bekannte Bedeutung haben. Aus (2) und (3) folgt

$$\left.\begin{aligned} \gamma^\mu \gamma_\nu + \gamma_\nu \gamma^\mu &= 2\,\delta_\nu^\mu \\ \text{und} \qquad \gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu &= 2\,g^{\mu\nu}. \end{aligned}\right\} \tag{4}$$

---

* P. A. M. Dirac, Proc. Roy. Soc. (A) **117**, 610 und **118**, 351, 1928; Friedrich Möglich, ZS. f. Phys. **48**, 852, 1928; J. v. Neumann, ebenda, S. 868.

** H. Tetrode, ZS. f. Phys. **49**, 858, 1928, als I zitiert.

*** Dirac spricht gelegentlich davon, daß die γ_μ Funktionen der Zeit wären; es sind dann aber die Volumenintegrale über $\bar{\psi}\,\gamma_\mu\,\psi$ gemeint.

Wir bilden den kovarianten Sechservektor mit den Komponenten

$$\alpha_{\mu\nu} = -\alpha_{\nu\mu} = \tfrac{1}{2}(\gamma_\mu \gamma_\nu - \gamma_\nu \gamma_\mu), \tag{5}$$

für deren Quadrate sich nach (2) ergibt:

$$\alpha_{\mu\nu}^2 = (\gamma_\mu \gamma_\nu - g_{\mu\nu})(g_{\mu\nu} - \gamma_\nu \gamma_\mu) = g_{\mu\nu}^2 - g_{\mu\mu} g_{\nu\nu}. \tag{6}$$

Diese sind also gewöhnliche Größen und gleich negativ genommenen Unterdeterminanten zweiter Ordnung der Determinante g des Fundamentaltensors.

Ferner bilden wir die skalare Dichte

$$\gamma = \frac{1}{4!} \sum_{\mu,\nu,\varrho,\sigma} \delta_{\mu\nu\varrho\sigma} \gamma_\mu \gamma_\nu \gamma_\varrho \gamma_\sigma \tag{7}$$

($\delta_{\mu\nu\varrho\sigma} = +1$ oder $= -1$, je nachdem $\mu\,\nu\,\varrho\,\sigma$ eine gerade oder eine ungerade Permutation der Indizes $1, 2, 3, 4$ ist; $\delta_{\mu\nu\varrho\sigma} = 0$, falls $\mu, \nu, \varrho, \sigma$ nicht sämtlich verschieden sind). Deren Quadrat γ^2 besteht aus einer Summe von Produkten von je acht Faktoren γ_μ, unter denen jedes γ_μ zweimal vorkommt. Wir können daher durch wiederholte Anwendung der Vertauschungsrelationen (2) erreichen, daß schließlich nur gleiche γ_μ nebeneinander vorkommen und dann γ_μ^2 durch $g_{\mu\mu}$ ersetzen. Es ergibt sich also für γ^2 eine Summe von Produkten von je vier Faktoren $g_{\mu\nu}$, unter deren acht Indizes jede der Zahlen 1, 2, 3, 4 zweimal vorkommen muß. Da γ^2 den Transformationscharakter des Quadrats einer skalaren Dichte hat, schließen wir, daß es bis auf einen Zahlenfaktor gleich der Determinante g der $g_{\mu\nu}$ sein muß. Der Faktor bestimmt sich für den Spezialfall (1) zu 1, und daher ist

$$\gamma^2 = g. \tag{8}$$

Die γ entsprechende kontravariante Bildung ist

$$\beta = \frac{1}{4!} \sum_{\mu,\nu,\varrho,\sigma} \delta_{\mu\nu\varrho\sigma} \gamma^\mu \gamma^\nu \gamma^\varrho \gamma^\sigma, \tag{9}$$

für deren Quadrat sich auf ebensolche Weise wie oben ergibt:

$$\beta^2 = \frac{1}{g}. \tag{10}$$

Ferner ist

$$\beta\gamma = \frac{1}{(4!)^2} \sum_{\mu,\nu,\varrho,\sigma,\alpha,\beta,\varkappa,\lambda} \delta_{\mu\nu\varrho\sigma} \delta_{\alpha\beta\varkappa\lambda} \gamma^\mu \gamma^\nu \gamma^\varrho \gamma^\sigma \gamma_\alpha \gamma_\beta \gamma_\varkappa \gamma_\lambda$$

$$= \frac{1}{(4!)^2} \sum_{\mu,\nu,\varrho,\sigma,\alpha,\beta,\varkappa,\lambda,a,b,k,l} g_{\alpha a} g_{\beta b} g_{\varkappa k} g_{\lambda l} \delta_{\mu\nu\varrho\sigma} \delta_{\alpha\beta\varkappa\lambda} \gamma^\mu \gamma^\nu \gamma^\varrho \gamma^\sigma \gamma^a \gamma^b \gamma^k \gamma^l.$$

Wegen

$$\sum_{\alpha,\beta,\varkappa,\lambda} \delta_{\alpha\beta\varkappa\lambda} g_{\alpha a} g_{\beta b} g_{\varkappa k} g_{\lambda l} = \delta_{abkl}\, g$$

ergibt sich

$$\beta\gamma = \beta^2 g = 1, \tag{11}$$

unter Beachtung von (10).

Schließlich bilden wir noch die kontravariante Vektordichte mit den Komponenten:

$$\beta^{\mu} = \frac{1}{3!} \sum_{\nu, \varrho, \sigma,} \delta_{\mu \nu \varrho \sigma} \gamma_{\nu} \gamma_{\varrho} \gamma_{\sigma}, \tag{12}$$

und finden nach einer ähnlichen Schlußweise, wie sie oben zur Ableitung von (8) angewandt wurde:

$$\beta^{\mu} \beta^{\nu} + \beta^{\nu} \beta^{\mu} = -2 g g^{\mu \nu}. \tag{13}$$

§ 2. Die allgemein-kovariante Verallgemeinerung der früheren Gleichungen I, (8) und (7) ist offenbar

$$\sum_{\mu} \gamma^{\mu} \left(\frac{h}{2 \pi} \frac{\partial \chi}{\partial x^{\mu}} + \frac{i e}{c} \varphi_{\mu} \chi \right) - m c \chi = 0 \tag{14}$$

und

$$\sum_{\mu} \left(-\frac{h}{2 \pi} \frac{\partial \omega}{\partial x^{\mu}} + \frac{i e}{c} \varphi_{\mu} \omega \right) \gamma^{\mu} - m c \omega = 0, \tag{15}$$

wo letztere Gleichung mit anderer Faktorenfolge geschrieben wurde.

Durch eine kanonische Transformation

$$\gamma_{\mu} \rightarrow S \gamma_{\mu} S^{-1}, \; \chi \rightarrow S \chi, \; \omega \rightarrow \omega S^{-1} \tag{16}$$

geht (14) über in

$$S \left\{ \sum_{\mu} \gamma^{\mu} \left(\frac{h}{2 \pi} \frac{\partial \chi}{\partial x^{\mu}} + \frac{h}{2 \pi} S^{-1} \frac{\partial S}{\partial x^{\mu}} \chi + \frac{i e}{c} \varphi_{\mu} \chi \right) - m c \chi \right\} = 0$$

oder, nach Multiplikation mit S^{-1}:

$$\sum_{\mu} \gamma^{\mu} \left(\frac{h}{2 \pi} \frac{\partial \chi}{\partial x^{\mu}} + \frac{h}{2 \pi} S^{-1} \frac{\partial S}{\partial x^{\mu}} \chi + \frac{i e}{c} \varphi_{\mu} \chi \right) - m c \chi = 0, \tag{17}$$

und Analoges gilt für (15). Würden wir also beliebige kanonische Transformationen zulassen, so wäre unser Gleichungssystem im allgemeinen für sie nicht mehr invariant. Im früheren Falle konstanter $g_{\mu \nu}$ erschien es sinnvoll, sich auf konstante Transformationsmatrizen S zu beschränken, wodurch (17) mit (14) identisch wird. Jetzt, wo die γ_{μ} Raumzeitfunktionen sind, wird man dasselbe auch für S annehmen müssen. Da jedoch die Invarianz von (14) gefordert werden muß, können wir nur solche S zulassen, für die gilt

$$\sum_{\mu} \gamma^{\mu} S^{-1} \frac{\partial S}{\partial x^{\mu}} = \sum_{\mu} \gamma^{\mu} \frac{\partial f}{\partial x^{\mu}}, \tag{18}$$

mit einer „gewöhnlichen" Größe f. Dann bedeutet (16) einfach Ersatz von χ durch χe^{f}, was nur eine Änderung der Bezeichnungsweise und daher physikalisch unwesentlich ist.

Wir werden jetzt zeigen, daß diese Bedingung erfüllt ist, falls die Rotation und die Divergenz von γ_μ durch (16) unverändert gelassen werden, in Formeln:

$$\frac{\partial \gamma_\mu}{\partial x^\nu} - \frac{\partial \gamma_\nu}{\partial x^\mu} = \frac{\partial (S \gamma_\mu S^{-1})}{\partial x^\nu} - \frac{\partial (S \gamma_\nu S^{-1})}{\partial x^\mu}$$

und

$$\sum_\mu \frac{\partial (\gamma^\mu \sqrt{-g})}{\partial x^\mu} = \sum_\mu \frac{\partial (S \gamma^\mu S^{-1} \sqrt{-g})}{\partial x^\mu},$$

oder*

$$\left.\begin{aligned} S^{-1} \frac{\partial S}{\partial x^\nu} \gamma_\mu - \gamma_\mu S^{-1} \frac{\partial S}{\partial x^\nu} &= S^{-1} \frac{\partial S}{\partial x^\mu} \gamma_\nu - \gamma_\nu S^{-1} \frac{\partial S}{\partial x^\mu} \\ \text{und} \qquad \sum_\mu S^{-1} \frac{\partial S}{\partial x^\mu} \gamma^\mu - \sum_\mu \gamma^\mu S^{-1} \frac{\partial S}{\partial x^\mu} &= 0. \end{aligned}\right\} \tag{19}$$

Zum Beweis genügt es offenbar, sich auf infinitesimale kanonische Transformationen zu beschränken. Für diese ist $S = 1 + \zeta$, mit unendlich kleiner Matrix ζ, und daher hat man in erster Näherung

$$S^{-1} \frac{\partial S}{\partial x^\mu} = (1 - \zeta) \frac{\partial \zeta}{\partial x^\mu} = \frac{\partial \zeta}{\partial x^\mu}. \tag{20}$$

Ferner wählen wir das Koordinatensystem für den betrachteten Punkt derart, daß die $g_{\mu\nu} = \delta_\mu^\nu$ werden, so daß zwischen Ko- und Kontravarianz nicht unterschieden zu werden braucht und die Gleichungen (1) gelten. Eine beliebige vierreihige Matrix ζ läßt sich darstellen als lineare Form der 16 linear-unabhängigen Matrizen 1, $\mathring{\gamma}_\mu$, $\mathring{\gamma}_\mu \mathring{\gamma}_\nu$, $\mathring{\gamma}_\mu \mathring{\gamma}_\nu \mathring{\gamma}_\varrho$, $\mathring{\gamma}_1 \mathring{\gamma}_2 \mathring{\gamma}_3 \mathring{\gamma}_4$ ($\mu \neq \nu \neq \varrho \neq \mu$), wo $\mathring{\gamma}_\mu$ den Wert von γ_μ in dem betrachteten Punkte bezeichnet, welcher Wert also als konstant zu betrachten ist und nicht differenziert werden muß. Dagegen sind die Koeffizienten der Form irgendwelche Funktionen der Koordinaten. Es ist also

$$\zeta = a + a_1 \mathring{\gamma}_1 + \cdots + a_{12} \mathring{\gamma}_1 \mathring{\gamma}_2 + \cdots + a_{123} \mathring{\gamma}_1 \mathring{\gamma}_2 \mathring{\gamma}_3 + \cdots + a_{1234} \mathring{\gamma}_1 \mathring{\gamma}_2 \mathring{\gamma}_3 \mathring{\gamma}_4,$$

und

$$\frac{\partial \zeta}{\partial x^\mu} = \frac{\partial a}{\partial x^\mu} + \frac{\partial a_1}{\partial x^\mu} \mathring{\gamma}_1 + \cdots + \frac{\partial a_{12}}{\partial x^\mu} \mathring{\gamma}_1 \mathring{\gamma}_2 + \cdots + \frac{\partial a_{123}}{\partial x^\mu} \mathring{\gamma}_1 \mathring{\gamma}_2 \mathring{\gamma}_3 + \cdots + \frac{\partial a_{1234}}{\partial x^\mu} \mathring{\gamma}_1 \mathring{\gamma}_2 \mathring{\gamma}_3 \mathring{\gamma}_4, \tag{21}$$

* Es ist von der Identität $\dfrac{\partial S^{-1}}{\partial x^\mu} = - S^{-1} \dfrac{\partial S}{\partial x^\mu} S^{-1}$ Gebrauch gemacht.

wo die $a_{\mu\nu}$, $a_{\mu\nu\varrho}$, $a_{\mu\nu\varrho\sigma}$ zweckmäßig als antisymmetrisch in den Indizes zu betrachten sind, da dasselbe nach (1) für die Produkte $\mathring{\gamma}_1 \mathring{\gamma}_2$, $\mathring{\gamma}_1 \mathring{\gamma}_2 \mathring{\gamma}_3$ usw. gilt, und es dann also auf die Reihenfolge der Faktoren nicht ankommt, insofern die Indizes an den a entsprechend angeordnet sind. Da nicht weiter differenziert werden soll, können wir jetzt den Index $\circ$ über den γ_μ in (21) wieder fortlassen und erhalten aus der ersten Gleichung (19) für $\mu = 1$, $\nu = 2$, unter Benutzung von (20) und (21) und indem wir mit Hilfe von (1) alle Produkte der γ_μ durch die 16 linear-unabhängigen 1, γ_1, $\dots \gamma_1 \gamma_2$ usw. ausdrücken:

$$\frac{\partial a_1}{\partial x^1} = -\frac{\partial a_2}{\partial x^2}, \qquad \frac{\partial a_3}{\partial x^1} = \frac{\partial a_4}{\partial x^1} = 0, \qquad \frac{\partial a_{12}}{\partial x^1} = \frac{\partial a_{12}}{\partial x^2} = 0,$$

$$\frac{\partial a_{32}}{\partial x^1} = \frac{\partial a_{31}}{\partial x^2}, \quad \frac{\partial a_{42}}{\partial x^2} = \frac{\partial a_{41}}{\partial x^1}, \quad \frac{\partial a_{341}}{\partial x^1} = -\frac{\partial a_{342}}{\partial x^2}, \quad \frac{\partial a_{1234}}{\partial x^1} = 0.$$

Da Entsprechendes auch für die anderen fünf Indexkombinationen μ, ν gelten muß, folgt

$$\frac{\partial a_1}{\partial x^1} = -\frac{\partial a_2}{\partial x^2} = \frac{\partial a_3}{\partial x^3} = -\frac{\partial a_1}{\partial x^1}$$

und also

$$\frac{\partial a_1}{\partial x^1} = 0,$$

ferner

$$\frac{\partial a_{32}}{\partial x^1} = \frac{\partial a_{31}}{\partial x^2} = -\frac{\partial a_{13}}{\partial x^2} = -\frac{\partial a_{12}}{\partial x^3} = \frac{\partial a_{21}}{\partial x^3} = \frac{\partial a_{23}}{\partial x^1} = -\frac{\partial a_{32}}{\partial x^1}$$

und also

$$\frac{\partial a_{32}}{\partial x^1} = 0.$$

Allgemein wird daher

$$\frac{\partial a_\mu}{\partial x^\varrho} = \frac{\partial a_{\mu\nu}}{\partial x^\varrho} = \frac{\partial a_{1234}}{\partial x^\varrho} = 0 \quad (\varrho \text{ beliebig}), \tag{22}$$

$$\frac{\partial a_{\mu\nu\varrho}}{\partial x^\varrho} = -\frac{\partial a_{\mu\nu\sigma}}{\partial x^\sigma} \quad (\mu, \nu, \varrho, \sigma \text{ verschieden}). \tag{23}$$

Nach (21) und (22) ist also zu setzen:

$$\frac{\partial \xi}{\partial x^\mu} = \frac{\partial a}{\partial x^\mu} + \frac{\partial a_{123}}{\partial x^\mu} \gamma_1 \gamma_2 \gamma_3 + \frac{\partial a_{234}}{\partial x^\mu} \gamma_2 \gamma_3 \gamma_4$$
$$+ \frac{\partial a_{341}}{\partial x^\mu} \gamma_3 \gamma_4 \gamma_1 + \frac{\partial a_{412}}{\partial x^\mu} \gamma_4 \gamma_1 \gamma_2. \tag{24}$$

Indem man dies, nach (20), in die zweite der Gleichungen (19) einsetzt, findet man, unter Beachtung von (1), daß diese Gleichung (19) erfüllt ist, wenn

$$\frac{\partial a_{123}}{\partial x_4} - \frac{\partial a_{234}}{\partial x_1} + \frac{\partial a_{341}}{\partial x_2} - \frac{\partial a_{412}}{\partial x_3} = 0 \tag{25}$$

ist. Aus (23), (24) und (25) ergibt sich schließlich

$$\sum_\mu \gamma_\mu \frac{\partial \xi}{\partial x^\mu} = \sum_\mu \gamma_\mu \frac{\partial a}{\partial x^\mu}. \tag{26}$$

Weil die beiden Seiten dieser Gleichung, nach Ersatz von γ_μ durch γ^μ, Skalare sind, gilt sie für jedes Koordinatensystem. Also ist (18) erfüllt.

§ 3. Wir wollen jetzt die Differentialgleichungen (14) und (15) aus einem einzigen Variationsprinzip

$$\delta \int\int\int\int \mathfrak{H}\, d x^1\, d x^2\, d x^3\, d x^4 = 0 \tag{27}$$

ableiten und setzen dazu

$$\mathfrak{H} = \omega \sqrt{-g} \left\{ \sum_\mu \gamma^\mu \left(\frac{h}{2\pi} \frac{\partial}{\partial x^\mu} + \frac{ie}{c} \varphi_\mu \right) - mc \right\} \chi, \tag{28}$$

was offenbar, wie erforderlich, eine skalare Dichte ist. Variation von ω liefert dann unmittelbar die Gleichung (14), Variation von χ dagegen ergibt, nach partieller Integration:

$$\sum_\mu \left(-\frac{h}{2\pi} \frac{\partial}{\partial x^\mu} + \frac{ie}{c} \varphi_\mu \right) \left(\omega \gamma^\mu \sqrt{-g} \right) - mc\omega \sqrt{-g} = 0, \tag{29}$$

was sich, nach Division durch $\sqrt{-g}$, von (15) durch den Term

$$-\frac{h}{2\pi} \frac{\omega}{\sqrt{-g}} \sum_\mu \frac{\partial \left(\gamma^\mu \sqrt{-g} \right)}{\partial x^\mu}$$

unterscheidet. Es steht nun nichts im Wege, aus den Gleichungen (14) und (29) bzw. aus dem Variationsprinzip (27), (28) ohne weiteres das Erhaltungsgesetz der Elektrizität sowie den Impuls-Energiesatz abzuleiten, genau so, wie es weiter unten geschehen wird. Weil aber (29) nicht wie (15) die „transponierte“ Gleichung zu (14) ist, würden wir nicht, wie es in I geschah, den richtigen Realitätscharakter unserer physikalischen Größen beweisen können*, und dieser würde ihnen dann auch im allgemeinen abgehen. Wir legen deshalb den γ_μ neben den algebraischen Bedingungen (2) noch die Differentialbedingung auf:

$$\frac{1}{\sqrt{-g}} \sum_\mu \frac{\partial \left(\gamma^\mu \sqrt{-g} \right)}{\partial x^\mu} = 0, \tag{30}$$

wodurch (29) mit (15) gleichbedeutend wird.

---

* Siehe den Zusatz am Schluß.

In § 2 hatten wir gefordert, daß die erlaubten kanonischen Transformationen die Divergenz und die Rotation von γ_μ ungeändert lassen sollten; darüber hinausgehend, fordern wir also jetzt das Verschwinden der Divergenz, welche Bedingung natürlich durch keine erlaubte kanonische Transformation verletzt werden darf. Man könnte daran denken, auch das Verschwinden der Rotation von γ_μ zu verlangen. Diese Bedingung würde jedoch, wie hier nicht näher ausgeführt werden soll, eine Beschränkung der zu fordernden freien Wählbarkeit des $g_{\mu\nu}$-Feldes bedeuten, indem aus ihr algebraische Beziehungen zwischen den Komponenten des Riemannschen Krümmungstensors resultieren würden. Wir schreiben also bloß die Bedingung (30) vor, die ja auch genügt, um die beiden Gleichungen (14) und (15) aus dem Variationsprinzip (27), (28) herleiten zu können.

§ 4. Durch Multiplikation von (14) mit $\omega\sqrt{-g}$, von (15) mit $\chi\sqrt{-g}$, und Subtraktion folgt, unter Beachtung von (30), das Erhaltungsgesetz der Elektrizität

$$\sum_\mu \frac{\partial \mathfrak{f}^\mu}{\partial x^\mu} = 0, \tag{31}$$

mit

$$\mathfrak{f}^\mu = -e\,\omega\,\gamma^\mu\sqrt{-g}\,\chi, \tag{32}$$

was übrigens auch aus (14) und (29), ohne (30), folgen würde.

Den Impuls-Energiesatz leiten wir jetzt ab durch eine infinitesimale Deformation des Koordinatensystems*. Bezeichnen wir mit δa die Variation einer Größe a für einen festgehaltenen, mit $\delta^* a$ diejenige für einen vom Koordinatensystem mitgenommenen Punkt, so gilt für eine skalare Dichte $\mathfrak{H}$

$$\delta\int \mathfrak{H}\,dx = \int \delta^*\mathfrak{H}\,dx \equiv 0 \tag{33}$$

für beliebige Variationen δx^μ, die nur an der Grenze des betrachteten Weltgebietes $\int dx = \int\int\int\int dx^1\,dx^2\,dx^3\,dx^4$ verschwinden sollen. Mit

$$\delta^*\gamma^\mu = \sum_\nu \frac{\partial\,\delta x^\mu}{\partial x^\nu}\gamma^\nu - \sum_\nu \delta x^\nu \frac{\partial \gamma^\mu}{\partial x^\nu},$$

$$\delta^*\varphi_\mu = -\sum_\nu \frac{\partial\,\delta x^\nu}{\partial x^\mu}\varphi_\nu - \sum_\nu \delta x^\nu \frac{\partial \varphi_\mu}{\partial x^\nu}$$

* Siehe z. B. W. Pauli, Relativitätstheorie, Nr. 23.

findet man für die Variation der Wirkungsgröße (27), nach (28), nachdem man die Ableitungen der δx^μ durch partielle Integration fortgeschafft hat:

$$\int \delta^* \mathfrak{H}\, d x = \sum_{\lambda, \mu} \Big[-\frac{\partial}{\partial x^\mu} \Big\{ \omega \sqrt{-g}\, \gamma^\mu \Big(\frac{h}{2\pi} \frac{\partial}{\partial x^\lambda} + \frac{i e}{c} \varphi_\lambda \Big) \chi \Big\}$$
$$- \omega \sqrt{-g} \frac{\partial \gamma^\mu}{\partial x^\lambda} \Big(\frac{h}{2\pi} \frac{\partial}{\partial x^\mu} + \frac{i e}{c} \varphi_\mu \Big) \chi + \frac{\partial}{\partial x^\mu} \Big(\omega \sqrt{-g}\, \gamma^\mu \frac{i e}{c} \varphi_\lambda \chi \Big)$$
$$- \omega \sqrt{-g}\, \gamma^\mu \frac{i e}{c} \frac{\partial \varphi_\mu}{\partial x^\lambda} \chi \Big] \delta x^\lambda. \tag{34}$$

Hierbei ist davon Gebrauch gemacht, daß der Koeffizient von $\delta^* (\omega \sqrt{-g})$ infolge von (14) verschwindet, ebenso wie derjenige von $\delta^* \chi$ infolge von (15) nach partieller Integration. Es muß also der Faktor von δx^λ in (34) Null sein, und unter Berücksichtigung von (31) und (32) findet man

$$\sum_\mu \frac{\partial}{\partial x^\mu} \Big\{ \frac{c}{i} \omega \sqrt{-g}\, \gamma^\mu \Big(\frac{h}{2\pi} \frac{\partial \chi}{\partial x^\lambda} + \frac{i e}{c} \varphi_\lambda \chi \Big) \Big\}$$
$$+ \sum_\mu \frac{c}{i} \omega \sqrt{-g} \frac{\partial \gamma^\mu}{\partial x^\lambda} \Big(\frac{h}{2\pi} \frac{\partial \chi}{\partial x^\mu} + \frac{i e}{c} \varphi_\mu \chi \Big)$$
$$+ \sum_\mu \mathfrak{f}^\mu \Big(\frac{\partial \varphi_\lambda}{\partial x^\mu} - \frac{\partial \varphi_\mu}{\partial x^\lambda} \Big) = 0. \tag{35}$$

Der letzte Term stellt die elektromagnetische Viererkraft auf die Volumeneinheit dar, die geschweifte Klammer im ersten Gliede wäre daher als der Impuls-Energietensor des Elektrons anzusehen, während dann das zweite Glied die vom Gravitationsfeld ausgeübte Viererkraftdichte wäre; nur haben die letztgenannten beiden Größen noch nicht den richtigen Realitätscharakter. Um dies zu erreichen, ersetzen wir $\mathfrak{H}$ durch die Größe $\mathfrak{H}'$, die daraus nach partieller Integration entsteht und die mit ebengleicher Berechtigung als Dichte der Wirkungsfunktion betrachtet werden kann:

$$\mathfrak{H}' = \Big\{ \sum_\mu \Big(-\frac{h}{2\pi} \frac{\partial \omega}{\partial x^\mu} + \frac{i e}{c} \varphi_\mu \omega \Big) \gamma^\mu - m c\, \omega \Big\} \sqrt{-g}\, \chi, \tag{36}$$

und erhalten

$$\sum_\mu \frac{\partial}{\partial x^\mu} \Big\{ \frac{c}{i} \Big(-\frac{h}{2\pi} \frac{\partial \omega}{\partial x^\lambda} + \frac{i e}{c} \varphi_\lambda \omega \Big) \sqrt{-g}\, \gamma^\mu \chi \Big\}$$
$$+ \sum_\mu \frac{c}{i} \Big(-\frac{h}{2\pi} \frac{\partial \omega}{\partial x^\mu} + \frac{i e}{c} \varphi_\mu \omega \Big) \sqrt{-g} \frac{\partial \gamma^\mu}{\partial x^\lambda} \chi$$
$$+ \sum_\mu \mathfrak{f}^\mu \Big(\frac{\partial \varphi_\lambda}{\partial x^\mu} - \frac{\partial \varphi_\mu}{\partial x^\lambda} \Big) = 0. \tag{37}$$

Durch Bildung des arithmetischen Mittels von (35) und (37) findet man schließlich

$$\sum_\mu \frac{\partial \mathfrak{T}_\lambda^\mu}{\partial x^\mu} + \sum_\mu \frac{c}{2}\left\{\omega \frac{\sqrt{-g}}{i} \frac{\partial \gamma^\mu}{\partial x^\lambda}\left(\frac{h}{2\pi}\frac{\partial \chi}{\partial x^\mu} + \frac{ie}{c}\varphi_\mu \chi\right)\right.$$

$$\left. + \left(-\frac{h}{2\pi}\frac{\partial \omega}{\partial x^\lambda} + \frac{ie}{c}\varphi_\mu \omega\right)\frac{\sqrt{-g}}{i}\frac{\partial \gamma^\mu}{\partial x^\lambda}\chi\right\} + \sum_\mu \mathfrak{f}^\mu \left(\frac{\partial \varphi_\lambda}{\partial x^\mu} - \frac{\partial \varphi_\mu}{\partial x^\lambda}\right) = 0, \tag{38}$$

mit

$$\mathfrak{T}_\lambda^\mu = \frac{c}{2}\omega \frac{\sqrt{-g}}{i}\gamma^\mu\left(\frac{h}{2\pi}\frac{\partial \chi}{\partial x^\lambda} + \frac{ie}{c}\varphi_\lambda \chi\right)$$

$$+ \frac{c}{2}\left(-\frac{h}{2\pi}\frac{\partial \omega}{\partial x^\lambda} + \frac{ie}{c}\varphi_\lambda \omega\right)\frac{\sqrt{-g}}{i}\gamma^\mu \chi, \tag{39}$$

welcher Ausdruck für $g_{\mu\nu} = \delta_\mu^\nu$ mit I, (14) übereinstimmt.

In der klassischen Theorie hat man statt des zweiten Gliedes von (38) einen Term

$$-\frac{1}{2}\sum_{\mu,\nu}\frac{\partial g_{\mu\nu}}{\partial x^\lambda}\mathfrak{T}^{\mu\nu} = \frac{1}{2}\sum_{\mu,\nu}\frac{\partial g^{\mu\nu}}{\partial x^\lambda}\mathfrak{T}_{\mu\nu}. \tag{40}$$

Man könnte versucht sein, durch Abänderung des Tensors $\mathfrak{T}$ aus (39), auch hier den entsprechenden Term auf diese Form zu bringen, indem man die allgemein-kovariante Verallgemeinerung des Flächentensors $T_{\lambda\mu} - T_{\mu\lambda}$ aus I, (16) heranzieht. Dies gelingt jedoch nicht, wie wir jetzt noch zeigen wollen.

Wir wählen ein Koordinatensystem, für das im betrachteten Punkte die $g_{\mu\nu} = \delta_\mu^\nu$ sind und das überdies dortselbst geodätisch ist. Es gelten hier die Gleichungen (1) für zwei unendlich benachbarte Punkte, so daß wir sie auch dann noch anwenden dürfen, wenn die γ_μ einmal differenziert werden sollen. Durch genau dieselben Überlegungen, die in I von (14) zu (16) führten, erhalten wir jetzt

$$\frac{2}{c}(T_{\lambda\mu} - T_{\mu\lambda})_{\lambda \neq \mu} = -\frac{h}{2\pi}\sum_{\substack{\nu \\ \lambda \neq \nu \neq \mu}}\left(\frac{\partial \omega}{\partial x^\nu}\gamma_\mu\gamma_\lambda\gamma_\nu\chi + \omega\gamma_\mu\gamma_\lambda\gamma_\nu\frac{\partial \chi}{\partial x^\nu}\right), \tag{41}$$

und für die Divergenz hiervon

$$\frac{2}{c}\sum_\mu \frac{\partial}{\partial x^\mu}(T_{\lambda\mu} - T_{\mu\lambda})$$

$$= -\frac{h}{2\pi}\sum_{\substack{\mu,\nu \\ \lambda \neq \nu \neq \mu \neq \lambda}}\left\{\frac{\partial \omega}{\partial x^\nu}\frac{\partial(\gamma_\mu\gamma_\lambda\gamma_\nu)}{\partial x^\mu}\chi + \omega\frac{\partial(\gamma_\mu\gamma_\lambda\gamma_\nu)}{\partial x^\mu}\frac{\partial \chi}{\partial x^\nu}\right\}. \tag{42}$$

Aus der Tatsache, daß hier die φ_μ, die in (39) vorkommen und ja nicht durch die besondere Wahl des Koordinatensystems wegtransformiert werden können, fortgefallen sind, schließt man, daß es nicht möglich sein kann, durch Zusatz eines antisymmetrischen Terms

$$\text{Konst.}\,(\mathfrak{T}^{\lambda\mu} - \mathfrak{T}^{\mu\lambda})$$

zu unserem ursprünglichen Tensor $\mathfrak{T}^{\lambda\mu}$ aus (39) die Viererkraftdichte des Gravitationsfeldes auf die Form (40) zu bringen.

Zusatz bei der Korrektur. Um die richtigen Realitätseigenschaften der $\mathfrak{s}^{\mu}$ und $\mathfrak{T}^{\mu}_{\lambda}$ nachweisen zu können, wählen wir das Koordinatensystem (x^1, x^2, x^3 reell, x^4 rein imaginär) so, daß

$$g\,g^{44} = 1, \quad g^{4k} = 0 \quad (k = 1, 2, 3) \tag{43}$$

wird. Dies muß im allgemeinen möglich sein, da wir vier Raumzeitfunktionen (die neuen Koordinaten als Funktionen der alten) beliebig wählen können und die linken Seiten in (43) nicht etwa sämtliche Komponenten eines Tensors, oder gar mehrerer Tensoren, bilden. Sodann setzen wir versuchsweise

$$\gamma^4 = \frac{1}{\sqrt{g}}\gamma^0_4, \quad \gamma^k = \sum_{m=1}^{3} a^k_m\,\gamma^0_m, \tag{44}$$

wo die γ^0_μ ($\mu = 1, 2, 3, 4$) konstante Hermitesche Matrizen bedeuten, die den Gleichungen (1) genügen, während die a^k_m ($k, m = 1, 2, 3$) „gewöhnliche" Raumzeitfunktionen sein sollen. Die zweite Gleichung (4) ist dann nach (43), (44) für $\mu = 4$, $\nu = 1, 2, 3, 4$ identisch erfüllt. Für $\mu, \nu = 1, 2, 3$ erhalten wir

$$g^{kl} = \sum_{m=1}^{3} a^k_m\, a^l_m \quad (k, l = 1, 2, 3). \tag{45}$$

Da die den g^{kl} ($k, l = 1, 2, 3$) entsprechende quadratische Differentialform (das räumliche Linienelement) positiv-definit ist, ist (45) durch reelle a^k_m erfüllbar, wie man unmittelbar erkennt, indem man für einen bestimmten Punkt durch Koordinatentransformation $g^{kl} = \delta^l_k$ macht. Also sind alle γ^μ ebenso wie die γ^0_μ hermitesch. Von den neun Funktionen a^k_m sind durch (45) nur sechs bestimmt. Die drei übrigen genügen gerade, die Bedingung (30) zu befriedigen, die für unseren Fall nach (44) durch die drei Gleichungen

$$\sum_{k=1}^{3} \frac{\partial\, a^k_m \sqrt{g}}{\partial x^k} = 0 \quad (m = 1, 2, 3) \tag{46}$$

ausgedrückt wird. Aus (15) folgt jetzt durch Multiplikation mit γ_4^0

$$\left(-\frac{h}{2\pi}\frac{\partial w}{\partial x^4}+\frac{ie}{c}\varphi_4 w\right)\frac{1}{\sqrt{g}}+\sum_{k=1}^{3}\left(-\frac{h}{2\pi}\frac{\partial w}{\partial x^k}+\frac{ie}{c}\varphi_k w\right)\gamma^k\gamma_4^0 - m c w \gamma_4^0 = 0, \tag{47}$$

und aus (14), wenn man $\chi = \gamma_4^0\psi$ setzt,

$$\left(\frac{h}{2\pi}\frac{\partial\psi}{\partial x^4}+\frac{ie}{c}\varphi_4\psi\right)\frac{1}{\sqrt{g}}+\sum_{k=1}^{3}\gamma^k\gamma_4^0\left(\frac{h}{2\pi}\frac{\partial\psi}{\partial x^k}+\frac{ie}{c}\varphi^k\psi\right)-mc\gamma_4^0\psi. \tag{48}$$

Da wegen (44) auch die $i\gamma^k\gamma_4^0$ ($k = 1, 2, 3$) hermitesch sind, sind (47) und (48) zueinander konjugiert-komplexe Gleichungen und es kann daher w zu $\psi = \gamma_4^0\chi$ konjugiert-komplex gewählt werden. Machen wir dann noch für einen bestimmten Punkt durch Koordinatentransformation $g^{\mu\nu} = \delta_\mu^\nu$, so kann man die richtigen Realitätsverhältnisse ebenso wie im spezial-relativistischen Falle dartun, da ja die Ableitungen der γ^μ in den $\mathfrak{s}^\mu$ und $\mathfrak{T}_\lambda^\mu$ nicht vorkommen. Führen wir jetzt eine beliebige Koordinatentransformation aus, so bleiben die Realitätsverhältnisse natürlich richtig.

Die kanonischen Transformationen aber sind nach § 2 auf eine enge Gruppe zu beschränken. Wir können daher für gegebene $g_{\mu\nu}$ nicht ein beliebiges, ihnen nach (2) entsprechendes System der γ_μ annehmen, sondern müssen es, so wie wir es jetzt getan haben, bestimmen, und haben dann nur noch die Wahl zwischen denjenigen, die hieraus durch Koordinatentransformationen sowie durch die nach § 2 zugelassenen kanonischen Transformationen entstehen. Auf gewisse formelle Fragen, die sich hieran knüpfen, können wir jetzt nicht eingehen.

Chapter 11
Vladimir Fock (1929): Geometrisierung der Diracschen Theorie des Elektrons

Vladimir Fock (1929), Geometrisierung der Diracschen Theorie des Elektrons. *Zeitschrift für Physik*, 57: 261–277.

Geometrisierung der Diracschen Theorie des Elektrons.

Von **V. Fock** in Leningrad.

(Eingegangen am 5. Juli 1929.)

Mit Hilfe des Begriffs der Parallelübertragung eines Halbvektors werden die Diracschen Gleichungen in allgemein invarianter Form geschrieben. Es werden der Energietensor gebildet und die makroskopischen sowie die quantenmechanischen Bewegungsgleichungen aufgestellt. Die ersteren haben die gewöhnliche Form: Divergenz des Energietensors gleich der Lorentzkraft, und die letzteren sind im wesentlichen mit denen der geodätischen Linie identisch. Das Auftreten des Viererpotentials φ_l neben den Riccikoeffizienten γ_{ikl} in der Formel für die Parallelübertragung gibt einerseits einen einfachen geometrischen Grund für das Auftreten des Ausdrucks $p_l - \frac{e}{c}\varphi_l$ in der Wellengleichung und zeigt andererseits, daß die Potentiale φ_l, abweichend von Einsteins Auffassung, einen selbständigen Platz im geometrischen Weltbild haben und nicht etwa Funktionen der γ_{ikl} sein müssen.

In einer Arbeit von D. Iwanenko und dem Verfasser* wurde die Vermutung ausgesprochen, daß die Diracschen Matrizen eine rein geometrische Bedeutung haben. In einer anderen Arbeit** dieser Autoren wurde der Begriff der Parallelverschiebung eines Halbvektors (d. h. eines Quadrupels von Größen, die sich wie die Diracschen ψ-Funktionen transformieren) aufgestellt.

In einer weiteren Notiz*** hat ferner der Verfasser diesen Begriff zur Aufstellung der allgemein-relativistischen Wellengleichung des Elektrons angewandt und die makroskopischen Bewegungsgleichungen in der Einsteinschen Form abgeleitet.

Die vorliegende Arbeit ist eine vervollständigte Wiedergabe der in der letztgenannten Notiz gegebenen Betrachtungen.

1. Die Transformationseigenschaften der Diracschen ψ-Funktionen wurden eingehend von F. Möglich**** und J. v. Neumann† studiert.

---

* V. Fock und D. Iwanenko, Über eine mögliche geometrische Deutung der relativistischen Quantentheorie. ZS. f. Phys. **54**, 798, 1929.

** Dieselben, Géométrie quantique linéaire et déplacement parallèle. C. R. **188**, 1470, 1929. Diese Arbeit wurde am 20. Mai 1929 in der physikalischen Konferenz in Charkow vorgetragen.

*** V. Fock, Sur les équations de Dirac dans la théorie de relativité générale. C. R. **189**, 25, 1929.

**** F. Möglich, Zur Quantentheorie des rotierenden Elektrons. ZS. f. Phys. **48**, 852, 1928.

† J. v. Neumann, Einige Bemerkungen zur Diracschen Theorie des relativistischen Drehelektrons. Ebenda S. 868.

Eine besonders einfache Form erhält das Transformationsgesetz, wenn man für die ersten drei der Diracschen α-Matrizen die Ausdrücke

$$\alpha_1 = \sigma_1, \quad \alpha_2 = \varrho_3 \sigma_2, \quad \alpha_3 = \sigma_3 \tag{1}$$

und als vierte Matrix eine der Matrizen

$$\alpha_4 = \varrho_2 \sigma_2, \quad \alpha_5 = \varrho_1 \sigma_2 \tag{1*}$$

wählt†, wo $\varrho_1, \varrho_2, \varrho_3$; $\sigma_1, \sigma_2, \sigma_3$ die von Dirac eingeführten vierreihigen Matrizen sind.

Dann entspricht nämlich einer allgemeinen Lorentztransformation die folgende Transformation der ψ-Funktionen:

$$\left.\begin{aligned} \psi_1' &= \alpha\,\psi_1 + \beta\,\psi_2; \quad & \psi_3' &= \bar{\alpha}\,\psi_3 + \bar{\beta}\,\psi_4, \\ \psi_2' &= \gamma\,\psi_1 + \delta\,\psi_2; \quad & \psi_4' &= \bar{\gamma}\,\psi_3 + \bar{\delta}\,\psi_4. \end{aligned}\right\} \tag{2}$$

Die komplexen Größen $\alpha, \beta, \gamma, \delta$ genügen der Bedingung

$$\alpha\delta - \beta\gamma = 1 \tag{3}$$

und gehen im Falle einer rein räumlichen Drehung in die gewöhnlichen Parameter von Cayley und Klein über.

Bezeichnet man mit α_0 die Einheitsmatrix, so bilden die Größen

$$A_i = \bar{\psi}\alpha_i\psi \quad (i = 0, 1, 2, 3) \tag{4}$$

die Komponenten eines Vierervektors, und die Größen

$$A_4 = \bar{\psi}\alpha_4\psi, \quad A_5 = \bar{\psi}\alpha_5\psi \tag{4*}$$

sind Invarianten. Diese Tatsache läßt sich in Formeln wie folgt ausdrücken. Bezeichnet man die Transformation (2) mit S:

$$\psi' = S\psi, \quad \bar{\psi}' = \bar{\psi}S^+, \tag{5}$$

wo S^+ die zu S adjungierte (transponiert-konjugierte) Matrix bezeichnet, so gelten die Gleichungen:

$$S^+\alpha_i S = \sum_{k=0}^{3} a_{ik}\alpha_k; \quad S^+\alpha_4 S = \alpha_4; \quad S^+\alpha_5 S = \alpha_5, \tag{6}$$

wo a_{ik} die Koeffizienten einer allgemeinen Lorentztransformation sind. Wegen

$$\bar{\psi}'\alpha_i\psi' = \bar{\psi}S^+\alpha_i S\psi$$

transformieren sich daher die Größen (4) und (4*) nach den Formeln

$$A_i' = \sum_{k=0}^{3} a_{ik}A_k; \quad A_4' = A_4; \quad A_5' = A_5, \tag{7}$$

d. h. wie ein Vierervektor bzw. wie Invarianten. Da die in den ψ quadratischen Größen A_i $(i = 0, 1, 2, 3)$ einen Vierervektor bilden, wollen wir die Größen ψ mit den Transformationseigenschaften (2) als „Halbvektor“ bezeichnen††.

† Siehe V. Fock, Über den Begriff der Geschwindigkeit in der Diracschen Theorie des Elektrons. Anhang. Ebenda **55**, 127, 1929.

†† Diese Bezeichnung wurde von Herrn L. Landau eingeführt.

Die expliziten Ausdrücke für die Größen A_i $(i = 0, 1, 2, 3, 4, 5)$ lauten:

$$\begin{aligned}
A_0 &= \bar{\psi}_1\psi_1 + \bar{\psi}_2\psi_2 + \bar{\psi}_3\psi_3 + \bar{\psi}_4\psi_4,\\
A_1 &= \bar{\psi}_1\psi_2 + \bar{\psi}_2\psi_1 + \bar{\psi}_3\psi_4 + \bar{\psi}_4\psi_3,\\
A_2 &= -i\bar{\psi}_1\psi_2 + i\bar{\psi}_2\psi_1 + i\bar{\psi}_3\psi_4 - i\bar{\psi}_4\psi_3,\\
A_3 &= \bar{\psi}_1\psi_1 - \bar{\psi}_2\psi_2 + \bar{\psi}_3\psi_3 - \bar{\psi}_4\psi_4,\\
A_4 &= -\bar{\psi}_1\psi_4 + \bar{\psi}_2\psi_3 + \bar{\psi}_3\psi_2 - \bar{\psi}_4\psi_1,\\
A_5 &= -i\bar{\psi}_1\psi_4 + i\bar{\psi}_2\psi_3 - i\bar{\psi}_3\psi_2 + i\bar{\psi}_4\psi_1.
\end{aligned}$$

An der Hand dieser Ausdrücke bestätigt man die folgende identische Beziehung zwischen den Größen A_i:

$$A_1^2 + A_2^2 + A_3^2 + A_4^2 + A_5^2 = A_0^2. \tag{8}$$

2. Wir haben die Transformationseigenschaften der ψ-Funktionen bei einer Lorentztransformation im Rahmen der speziellen Relativitätstheorie betrachtet. Stellen wir uns auf den Standpunkt der allgemeinen Relativitätstheorie, so müssen wir, um den Begriff des Halbvektors einführen zu können, in jedem Raumzeitpunkt ein orthogonales (genauer pseudoorthogonales) Bezugssystem haben. Zu diesem Zwecke führen wir ein Netz von vier orthogonalen Kurvenkongruenzen ein und bezeichnen nach Einstein die Richtungen dieser Kongruenzen als „Beine". Die Betrachtungen des vorigen Paragraphen bleiben dann auch für den Fall der allgemeinen Relativitätstheorie gültig, wenn wir unter A_i die Komponenten eines Vektors nach den Beinen verstehen.

Wir numerieren die Beine mit lateinischen und die Koordinaten mit griechischen Indizes, die überall die Werte 0, 1, 2, 3 durchlaufen; bei der Summierung nach den lateinischen Indizes wird das Summenzeichen explizite angegeben, bei der Summierung nach den griechischen Indizes dagegen unterdrückt. Die Parameter der Kurvenkongruenzen bezeichnen wir mit h_k^α und die Momente mit $h_{k,\alpha}$. Da wir es mit einer indefiniten Metrik zu tun haben, führen wir mit Eisenhart* die Größen $e_1 = e_2 = e_3 = -1$; $e_0 = +1$ ein. Die Komponenten eines Vektors nach den Koordinatenrichtungen (A_σ) und nach den Beinen (A'_k)** drücken sich dann die einen durch die anderen wie folgt aus:

$$A'_k = A_\sigma h_k^\sigma; \qquad A_\sigma = \sum_k e_k A'_k h_{k,\sigma}. \tag{9}$$

* L. Eisenhart, Riemannian Geometry. Princeton 1926. — Siehe auch die vortreffliche Zusammenstellung der wichtigsten Formeln und Tatsachen in der Arbeit von T. Levi-Civita, „Vereinfachte Herstellung der Einsteinschen einheitlichen Feldgleichungen". Berl. Ber. 1929, S. 3.

** Die Bein- und die Koordinatenkomponenten werden im folgenden öfters mit einem und demselben Buchstaben bezeichnet; um Verwechslungen zu vermeiden, werden dabei die ersteren mit einem Strich versehen.

Bezeichnen wir mit $d s_k$ die Beinkomponenten einer infinitesimalen Verschiebung, so folgt aus der Formel

$$\delta A_\alpha = \Gamma^\beta_{\alpha\sigma} A_\beta d x^\sigma; \quad \Gamma^\beta_{\alpha\sigma} = \left\{ {\alpha\sigma \atop \beta} \right\} \tag{10}$$

für die Änderung der Komponenten eines Vektors bei einer Parallelverschiebung der folgende Ausdruck für die Änderung seiner Beinkomponenten:

$$\delta A'_i = \sum_{k\,l} e_k e_l \gamma_{i\,k\,l} A'_k d s_l, \tag{11}$$

wo $\gamma_{i\,k\,l}$ die von Ricci eingeführten Rotationskoeffizienten sind:

$$\gamma_{i\,k\,l} = (\nabla_\sigma h^\beta_i) h_{k,\beta} h^\sigma_l = (\nabla_\sigma h_{i,\beta}) h^\beta_k h^\sigma_l. \tag{12}$$

Dabei bezeichnet ∇_σ die kovariante Ableitung nach x^σ.

3. Wir wollen nun die Änderung der Komponenten eines Halbvektors ψ bei einer infinitesimalen Parallelverschiebung betrachten. Für diese Änderung machen wir den Ansatz

$$\delta \psi = \sum_l e_l C_l d s_l \psi. \tag{13}$$

Die C_l sind Matrizen mit den Elementen $(C_l)_{m\,n}$, und unter $C_l \psi$ verstehen wir vier Funktionen, deren m-te durch die Formel

$$(C_l \psi)_m = \sum_{n=1}^{4} (C_l)_{m\,n} \psi_n$$

gegeben wird.

Die zu (13) konjugiert-komplexe Gleichung lautet:

$$\delta \overline{\psi} = \bar{\psi} \sum_l e_l C^+_l d s_l, \tag{13*}$$

wo C^+_l die adjungierte Matrix bezeichnet. Nun ist durch das Gesetz (13) der Parallelverschiebung eines Halbvektors dasjenige eines Vektors bereits bestimmt; wir müssen nämlich haben:

$$\begin{aligned} \delta A'_i &= \delta (\overline{\psi} \alpha_i \psi) = \delta \overline{\psi} \alpha_i \psi + \overline{\psi} \alpha_i \delta \psi \\ &= \overline{\psi} \sum_l e_l (C^+_l \alpha_i + \alpha_i C_l) d s_l \psi. \end{aligned} \tag{14}$$

Soll diese Änderung mit der durch (11) gegebenen übereinstimmen, so müssen die C_l den Bedingungen

$$C^+_l \alpha_i + \alpha_i C_l = \sum_k e_k \alpha_k \gamma_{i\,k\,l} \tag{15}$$

genügen. Da ferner $A'_4 = \overline{\psi} \alpha_4 \psi$ und $A'_5 = \overline{\psi} \alpha_5 \psi$ Invarianten sind, so muß

$$\delta A'_4 = \overline{\psi} \sum_l e_l (C^+_l \alpha_4 + \alpha_4 C_l) d s_l \psi \tag{16}$$

und ebenso $\delta A'_5$ verschwinden, woraus die weiteren Bedingungen

$$C^+_l \alpha_4 + \alpha_4 C_l = 0; \quad C^+_l \alpha_5 + \alpha_5 C_l = 0 \tag{17}$$

folgen.

Man überzeugt sich unmittelbar, daß die allgemeine Lösung der Gleichungen (15) und (17) durch die Formel

$$C_l = \tfrac{1}{4} \sum_{m\,k} \alpha_m \alpha_k e_k \gamma_{mkl} + i\,\Phi_l' \tag{18}$$

gegeben wird, wo Φ_l' hermitische Matrizen sind, die mit allen α_i, sowie mit α_4 und α_5 vertauschbar sein müssen. Bleibt man im Gebiet der vierreihigen Matrizen, so folgt aus der Vertauschbarkeit mit allen α-Matrizen die Proportionalität mit der Einheitsmatrix. Betrachtet man dagegen Matrizen mit mehr als vier Reihen*, so ist der Fall nicht ausgeschlossen, daß die Φ_l' nicht der Einheitsmatrix proportional sind. Wir wollen bei den vierreihigen Matrizen bleiben und die Φ_l' als reelle Zahlen betrachten.

Zu beachten ist, daß die C_l die Matrizen α_4 und α_5 nicht enthalten, so daß sie die ersten zwei ψ-Funktionen unter sich und die letzten zwei unter sich transformieren. Wegen der Formeln (2) war das auch a priori zu erwarten.

4. Nachdem wir den Begriff der Parallelverschiebung eines Halbvektors aufgestellt haben, können wir denjenigen der kovarianten Ableitung $D_l'\psi$ eines Halbvektors ψ nach der Beinrichtung l durch die Formel

$$D_l'\psi = \frac{\partial \psi}{\partial s_l} - C_l\psi \tag{19}$$

definieren, wo $\dfrac{\partial \psi}{\partial s_l} = h_l^\sigma \cdot \dfrac{\partial \psi}{\partial x^\sigma}$ die Ableitung nach der Richtung des l-ten Beines bezeichnet. Die kovariante Ableitung eines Halbvektors nach der Koordinate x^σ bezeichnen wir mit

$$D_\sigma \psi = \frac{\partial \psi}{\partial x^\sigma} - \Gamma_\sigma \psi, \tag{19*}$$

wo zur Abkürzung

$$\Gamma_\sigma = \sum_k e_k h_{k,\,\sigma} C_k \tag{20}$$

gesetzt ist.

Betrachtet man für einen Augenblick den Raum als pseudoeuklidisch und setzt man γ_{ikl} gleich Null, so wird der Ausdruck (20) für D_l' gleich

$$D_l'\psi = \frac{\partial \psi}{\partial s_l} - i\,\Phi_l'\psi.$$

Das ist aber gerade der Ausdruck, der in der Diracschen Gleichung vorkommt, wenn man unter Φ_l' die Größe

$$\Phi_l' = \frac{2\pi e}{hc}\varphi_l' \tag{21}$$

* Solche Matrizen könnten vielleicht bei gewissen Verallgemeinerungen der Diracschen Gleichung, z. B. auf das Zweikörperproblem, vorkommen.

versteht, wo φ'_l die Beinkomponente des Vektorpotentials bezeichnet. Wir wollen im folgenden an dieser physikalischen Interpretation der geometrischen Größen Φ'_l festhalten. Wir haben somit für das Auftreten des Vektorpotentials in der Diracschen Gleichung eine geometrische Deutung gewonnen; und zwar ist diese Deutung derart, daß das Potential auch dann von Null verschieden sein kann, wenn die die Größen γ_{ikl} enthaltenden Gravitationsglieder verschwinden.

Wenden wir uns nun der Formel (13) für $\delta\psi$ zu, so sehen wir, daß dort gerade die Weylsche lineare Differentialform

$$\sum_l e_l \varphi'_l d s_l = \varphi_\sigma d x^\sigma$$

auftritt, in Übereinstimmung mit der von Weyl ausgesprochenen Vermutung*. Das Auftreten der Weylschen Differentialform im Gesetz der Parallelverschiebung eines Halbvektors steht in enger Beziehung mit der vom Verfasser** und auch von Weyl (l. c.) bemerkten Tatsache, daß die Addition eines Gradienten zum Viererpotential der Multiplikation der ψ-Funktion mit einem Faktor vom absoluten Betrag 1 entspricht. Diese Tatsache wurde von Weyl als „Prinzip der Eichinvarianz" bezeichnet.

5. Der Begriff der kovarianten Ableitung eines Halbvektors ermöglicht es, die Diracsche Wellengleichung für das Elektron in der allgemeinen Relativitätstheorie aufzustellen. Zu diesem Zwecke betrachten wir den Operator

$$F\psi = \frac{h}{2\pi i}\sum_k e_k \alpha_k \left(\frac{\partial \psi}{\partial s_k} - C_k \psi\right) - m c \alpha_4 \psi. \tag{22}$$

Wir wollen zeigen, daß er selbstadjungiert ist***. Um dieses einzusehen, gehen wir von den Beinen zu den Koordinaten über und führen die Matrizen

$$\gamma^\sigma = \sum_k e_k \alpha_k h_k^\sigma \tag{23}$$

und die durch (20) definierten Matrizen Γ_σ ein. Aus den Gleichungen (15) folgen analoge Beziehungen für die soeben eingeführten Matrizen

$$\Gamma_\alpha^+ \gamma^\sigma + \gamma^\sigma \Gamma_\alpha = -\nabla_\alpha \gamma^\sigma. \tag{24}$$

Diese Formel läßt sich leicht beweisen, indem man auf die Definition (12) der γ_{ikl} zurückgeht.

---

* H. Weyl, Gruppentheorie und Quantenmechanik, § 19, S. 88. Leipzig 1928.

** V. Fock, Über die invariante Form der Wellen- und der Bewegungsgleichungen für einen geladenen Massenpunkt. ZS. f. Phys. **39**, 226, 1926.

*** Das Wort „selbstadjungiert" verstehen wir hier in einem etwas erweiterten Sinne. Wir meinen nämlich darunter, daß der Ausdruck $\overline{\psi} F\psi - \overline{(F\psi)}\psi$ in Form einer (im allgemeinen vierdimensionalen) Divergenz geschrieben werden kann.

Durch die Koordinaten ausgedrückt, lautet der Operator F:

$$F\psi = \frac{h}{2\pi i}\gamma^\sigma\left(\frac{\partial\psi}{\partial x^\sigma} - \Gamma_\sigma\psi\right) - mc\alpha_4\psi. \tag{25}$$

Unter Beachtung von (24) beweist man leicht die Identität

$$\overline{\psi}F\psi - \overline{(F\psi)}\psi = \frac{h}{2\pi i}\frac{1}{\sqrt{g}}\frac{\partial}{\partial x^\sigma}(\overline{\psi}\sqrt{g}\gamma^\sigma\psi), \tag{26}$$

wo g den Absolutwert der Determinante $\|g_{\varrho\sigma}\|$ bezeichnet. Diese Identität drückt die Tatsache aus, daß der Operator F selbstadjungiert ist. Diese Tatsache gestattet es, für die Diracsche Gleichung in der allgemeinen Relativitätstheorie den Ansatz

$$F\psi = 0 \tag{27}$$

zu machen. Genügt ψ dieser Gleichung, so folgt aus der Identität (26), daß die Divergenz des Stromvektors

$$S^\varrho = \overline{\psi}\gamma^\varrho\psi, \tag{28}$$

der wegen des hermitischen Charakters der Matrizen γ^ϱ offenbar reell ist, verschwindet:

$$\frac{1}{\sqrt{g}}\frac{\partial}{\partial x^\sigma}(\sqrt{g}S^\sigma) = 0. \tag{29}$$

Es läßt sich leicht beweisen †, daß der Ansatz (25) und (27) für die Diracsche Gleichung nicht nur in bezug auf die Koordinatenwahl, sondern auch in bezug auf die Wahl der orthogonalen Kurvenkongruenzen invariant (genauer: kovariant) ist.

Zum Beweis bemerken wir zunächst, daß die Γ_σ eindeutig und übereinstimmend mit der früheren Definition (18), (20), (21) durch die Gleichungen

$$\left.\begin{aligned} \Gamma_\alpha^+\gamma^\sigma + \gamma^\sigma\Gamma_\alpha &= -\nabla_\alpha\gamma^\sigma \\ \frac{1}{4}\,\mathrm{Spur}\,\Gamma_\sigma &= \frac{2\pi i e}{hc}\varphi_\sigma \end{aligned}\right\} \tag{30}$$

definiert werden können. Führt man jetzt irgend ein neues Netz von Kurvenkongruenzen ein und bezeichnet man die auf dieses Netz bezüglichen Größen mit einem Stern, so sind die neuen Γ_σ^*-Lösungen der analogen Gleichungen

$$\left.\begin{aligned} \Gamma_\alpha^{*+}\gamma^\sigma + \gamma^{*\sigma}\Gamma_\alpha^* &= -\nabla_\alpha\gamma^{*\sigma} \\ \frac{1}{4}\,\mathrm{Spur}\,\Gamma_\sigma^* &= \frac{2\pi i e}{hc}\varphi_\sigma. \end{aligned}\right\} \tag{30*}$$

† Dieser Absatz (bis Ende des § 5) ist bei der Korrektur zugefügt.

Der Übergang zu neuen Beinrichtungen hat aber in jedem Raumzeitpunkt den Charakter einer lokalen Lorentztransformation. Folglich sind die neuen Komponenten des Halbvektors ψ^* und die neuen Matrizen $\gamma^{*\sigma}$ mit den alten ψ und γ^σ durch Relationen von der Form

$$\psi^* = S\psi; \qquad \gamma^\sigma = S^+ \gamma^{*\sigma} S \tag{31}$$

[vgl. Formeln (5) und (6)] verknüpft, wo S eine Matrix von der Form

$$S = \begin{Bmatrix} \alpha\,\beta\,0\,0 \\ \gamma\,\delta\,0\,0 \\ 0\,0\,\bar{\alpha}\,\bar{\beta} \\ 0\,0\,\bar{\gamma}\,\bar{\delta} \end{Bmatrix}, \qquad \alpha\delta - \beta\gamma = 1$$

mit variablen Elementen bezeichnet.

Das Transformationsgesetz für die Koeffizienten Γ_σ der Parallelübertragung lautet aber:

$$\Gamma_\sigma^* = S\Gamma_\sigma S^{-1} + \frac{\partial S}{\partial x^\sigma} S^{-1}, \tag{32}$$

denn dieser Ausdruck ist die eindeutige Lösung von (30*)†.

Es gilt ferner die Beziehung

$$\frac{\partial \psi^*}{\partial x^\sigma} - \Gamma_\sigma \psi^* = S\left(\frac{\partial \psi}{\partial x^\sigma} - \Gamma_\sigma \psi\right). \tag{33}$$

Bezeichnet man mit $F^*\psi^*$ den zu (25) analogen Ausdruck, den man bekommt, wenn man dort γ^σ, Γ_σ und ψ mit einem Stern versieht, so folgt aus (31) und (33)

$$F\psi = S^+ F^* \psi^*. \tag{34}$$

Die Gleichung $F\psi = 0$ ist also mit $F^*\psi^* = 0$ gleichbedeutend, was zu beweisen war.

6. In diesem Paragraphen wollen wir den Operator F in einer anderen Form darstellen, indem wir die in der Formel (22) auftretende Summe $\sum\limits_k e_k \alpha_k C_k$ berechnen.

Um das Resultat in übersichtlicher Form darstellen zu können, verfahren wir folgendermaßen. Wir führen die Größen ε_{ijkl} ein, welche verschwinden sollen, wenn unter den Indizes $ijkl$ zwei gleiche vorkommen, und im Falle verschiedener Indizes gleich $+1$ oder -1 sind, je nachdem die Zahlenfolge $ijkl$ aus 0 1 2 3 durch eine gerade oder eine ungerade Permutation hervorgeht. Mit Hilfe dieser Größen bilden wir den „Beinvektor"

$$f_i = \frac{1}{2} \sum_{jkl} e_j e_k e_l \varepsilon_{ijkl} \gamma_{jkl} \tag{35}$$

† Es gilt: Spur $\frac{\partial S}{\partial x^\sigma} S^{-1} = 0$.

mit den Komponenten

$$\left.\begin{aligned} f_0 &= -e_0(\gamma_{123} + \gamma_{231} + \gamma_{312}), \\ f_1 &= -e_1(\gamma_{203} + \gamma_{032} + \gamma_{320}), \\ f_2 &= -e_2(\gamma_{301} + \gamma_{013} + \gamma_{130}), \\ f_3 &= -e_3(\gamma_{102} + \gamma_{021} + \gamma_{210}). \end{aligned}\right\} \tag{35*}$$

Beachten wir die Identitäten

$$\left.\begin{aligned} \alpha_1\alpha_2\alpha_3 &= i\varrho_3\alpha_0, \\ \alpha_2\alpha_3 &= i\varrho_3\alpha_1, \\ \alpha_3\alpha_1 &= i\varrho_3\alpha_2, \\ \alpha_1\alpha_2 &= i\varrho_3\alpha_3, \end{aligned}\right\} \tag{1**}$$

die aus der Definition (1) der Matrizen α_i hervorgehen, so können wir die Summe $\sum\limits_k e_k\alpha_k C_k$ in der Form

$$\sum_l e_l\alpha_l C_l = \sum_l e_l\alpha_l\left(i\Phi_l' - \frac{1}{2}\sum_j e_j\gamma_{jlj} - \frac{i}{2}\varrho_3 f_l\right) \tag{*}$$

schreiben. Wir bezeichnen

$$k_i = -\sum_j e_j\gamma_{jij} = \frac{1}{\sqrt{g}}\frac{\partial}{\partial x^\sigma}(\sqrt{g}\,h_i^\sigma) \tag{36}$$

und führen den Ausdruck (*) in (22) ein. Wir erhalten dann

$$\begin{aligned} F\psi = \sum_j e_j\alpha_j\left(\frac{h}{2\pi i}\frac{\partial\psi}{\partial s_j} - \frac{e}{c}\varphi_j'\psi + \frac{h}{4\pi i}k_j\psi\right) \\ + \frac{h}{4\pi}\varrho_3\sum_j e_j\alpha_j f_j\psi - mc\alpha_4\psi. \end{aligned} \tag{22*}$$

Wir bemerken, daß in diesem Ausdruck die erste und die zweite Summe einzeln selbstadjungierte Operatoren sind.

Falls alle Kongruenzen Normalenkongruenzen sind, verschwindet der „Beinvektor“ f_i, da jedes Riccisymbol γ_{ikl} mit drei verschiedenen Indizes einzeln verschwindet; ferner können wir dann die Hyperflächen, deren Normalen die Kurvenkongruenzen geben, als Koordinatenflächen wählen. Wir haben dann

$$ds^2 = \sum_j e_j H_j^2\,dx_j^2; \qquad \sqrt{g} = H_0 H_1 H_2 H_3; \tag{37}$$

$$h_i^i = \frac{1}{H_i}; \qquad h_{i,i} = e_i H_i; \qquad f_i = 0, \tag{37*}$$

während alle Parameter $h_{i,\sigma}$ und h_i^σ mit verschiedenen Indizes verschwinden. Der Ausdruck für den Operator F lautet dann

$$F\psi = \sum_j e_j \alpha_j \frac{1}{H_j}\left(\frac{h}{2\pi i}\frac{\partial\psi}{\partial x_j} - \frac{e}{c}\varphi_j\psi + \frac{h}{4\pi i}\frac{\partial}{\partial x_j}\left(\lg\frac{\sqrt{g}}{H_j}\right)\psi\right) - mc\alpha_4\psi. \tag{38}$$

Diese Formel erlaubt es, die Diracsche Gleichung in beliebigen krummlinigen orthogonalen Koordinaten sofort hinzuschreiben. Dabei ist folgendes zu beachten. Schreibt man z. B. im Falle eines gewöhnlichen Euklidischen Raumes die Gleichung (38) einmal in kartesischen, ein anderes Mal in krummlinigen Koordinaten, so sind die in (38) in beiden Fällen auftretenden ψ-Funktionen nicht identisch, sondern miteinander durch eine Transformation von der Form (2) mit variablen Koeffizienten α, β, γ, δ verknüpft. Diesen Umstand muß man bei der Aufstellung der Eindeutigkeitsforderungen für die ψ-Funktionen im Auge behalten.

Zum Schlusse dieses Paragraphen sei hier bemerkt, daß es bekanntlich in einem allgemeinen Riemannschen Raume nicht immer möglich ist, alle Kurvenkongruenzen als Normalenkongruenzen zu wählen. Das ist jedoch in den wichtigen Spezialfällen eines statischen Gravitationsfeldes mit zentraler und axialer Symmetrie jedenfalls möglich, wie es die von Schwarzschild und Levi-Civita gefundenen Lösungen der Einsteinschen Gleichungen gezeigt haben.

7. Wir wollen jetzt versuchen, den Energietensor zu finden. Zu diesem Zwecke betrachten wir den Tensor

$$A^{\sigma}_{\cdot\alpha} = \bar{\psi}\gamma^\sigma\left(\frac{\partial\psi}{\partial x^\alpha} - \Gamma_\alpha\psi\right) = \bar{\psi}\gamma^\sigma D_\alpha\psi \tag{39}$$

und berechnen seine Divergenz †.

Wir schreiben die Diracsche Gleichung mit ihrer konjugiert komplexen in der Form

$$\gamma^\sigma\left(\frac{\partial\psi}{\partial x^\sigma} - \Gamma_\sigma\psi\right) - \frac{2\pi i}{h}mc\alpha_4\psi = 0, \tag{40}$$

$$\left(\frac{\partial\bar{\psi}}{\partial x^\sigma} - \bar{\psi}\Gamma_\sigma^+\right)\gamma^\sigma + \frac{2\pi i}{h}mc\bar{\psi}\alpha_4 = 0. \tag{40*}$$

---

† Die Resultate dieses Paragraphen könnte man auch in eleganterer Form durch Betrachtung einer infinitesimalen Koordinatentransformation ableiten (vgl. H. Tetrode, Allgemein-relativistische Quantentheorie des Elektrons, ZS. f. Phys. **50**, 336, 1928). Wir ziehen es jedoch vor, in mehr elementarer Weise vorzugehen.

Wir differentiieren (40) nach x^α und multiplizieren links mit $\overline{\psi}$; die Gleichung (40*) multiplizieren wir rechts mit $\frac{\partial \psi}{\partial x^\alpha}$ und addieren die Resultate. Beachten wir die aus (24) folgende Formel

$$\Gamma_\sigma^+ \gamma^\sigma + \gamma^\sigma \Gamma_\sigma = -\frac{1}{\sqrt{g}} \frac{\partial \sqrt{g}\, \gamma^\sigma}{\partial x^\sigma}, \tag{41}$$

so können wir die Summe in der Form

$$\frac{1}{\sqrt{g}} \frac{\partial}{\partial x^\sigma}\left(\overline{\psi} \sqrt{g}\, \gamma^\sigma \frac{\partial \psi}{\partial x^\alpha}\right) + \overline{\psi} \frac{\partial \gamma^\sigma}{\partial x^\alpha} D_\sigma \psi - \overline{\psi} \gamma^\sigma \frac{\partial \Gamma_\sigma}{\partial x^\alpha} \psi = 0 \tag{42}$$

schreiben. Ferner multiplizieren wir (40) links mit $-\overline{\psi}\Gamma_\alpha^+$, (40*) rechts mit $\Gamma_\alpha \psi$ und addieren; in der Summe ersetzen wir $\Gamma_\alpha^+ \gamma^\sigma$ und $\Gamma_\sigma^+ \gamma^\sigma$ durch ihre Ausdrücke aus (24) und (41). Wir bekommen auf diese Weise

$$\begin{aligned}\frac{1}{\sqrt{g}} \frac{\partial}{\partial x^\sigma} &(\overline{\psi} \sqrt{g}\, \gamma^\sigma \Gamma_\alpha \psi) + \overline{\psi} (\nabla_\alpha \gamma^\sigma) D_\sigma \psi \\ &- \overline{\psi} \gamma^\sigma \frac{\partial \Gamma_\alpha}{\partial x^\sigma} \psi + \overline{\psi} \gamma^\sigma (\Gamma_\sigma \Gamma_\alpha - \Gamma_\alpha \Gamma_\sigma) \psi = 0.\end{aligned} \tag{43}$$

Ersetzen wir hier $\nabla_\alpha \gamma^\sigma$ durch

$$\nabla_\alpha \gamma^\sigma = \frac{\partial \gamma^\sigma}{\partial x^\alpha} + \Gamma^\sigma_{\alpha\varrho} \gamma^\varrho,$$

so ergibt die Subtraktion von (43) aus (42)

$$\frac{1}{\sqrt{g}} \frac{\partial}{\partial x^\sigma} (\sqrt{g}\, A^{\sigma}_{\cdot\alpha}) - \Gamma^\sigma_{\alpha\varrho} A^{\varrho}_{\cdot\sigma} = \overline{\psi} \gamma^\sigma D_{\sigma\alpha} \psi, \tag{44}$$

wo zur Abkürzung

$$D_{\sigma\alpha} = \frac{\partial \Gamma_\sigma}{\partial x^\alpha} - \frac{\partial \Gamma_\alpha}{\partial x^\sigma} + \Gamma_\sigma \Gamma_\alpha - \Gamma_\alpha \Gamma_\sigma \tag{45}$$

gesetzt ist.

Wir müssen nun die Matrix $D_{\sigma\alpha}$ berechnen. Wir haben

$$D_{\sigma\alpha} = D_\sigma D_\alpha - D_\alpha D_\sigma = \sum_{kl} e_k e_l h_{k,\sigma} h_{l,\alpha} D'_{kl}, \tag{46}$$

wo

$$D'_{kl} = D'_k D'_l - D'_l D'_k + \sum_m e_m (\gamma_{mlk} - \gamma_{mkl}) D'_m \tag{47}$$

gesetzt ist. Der Operator (47) ist gleich

$$D'_{kl} = \frac{1}{4} \sum_{ij} \alpha_i \alpha_j e_j \gamma_{ijkl} + \frac{2\pi i e}{hc} M'_{kl}, \tag{48}$$

wo γ_{ijkl} die Beinkomponenten des Riemannschen Tensors bezeichnen:

$$\begin{aligned}\gamma_{ijkl} &= \frac{\partial \gamma_{ijk}}{\partial s_l} - \frac{\partial \gamma_{ijl}}{\partial s_k} \\ &+ \sum_m e_m [\gamma_{ijm} (\gamma_{mkl} - \gamma_{mlk}) + \gamma_{mil} \gamma_{mjk} - \gamma_{mik} \gamma_{mjl}],\end{aligned} \tag{49}$$

und der schiefsymmetrische Tensor M'_{kl}:

$$M'_{kl} = \frac{\partial \varphi'_k}{\partial s_l} - \frac{\partial \varphi'_l}{\partial s_k} + \sum_m e_m (\gamma_{mlk} - \gamma_{mkl})\, \varphi'_m, \tag{50}$$

das elektromagnetische Feld darstellt.

Die Matrix $\gamma^\sigma D_{\sigma\alpha}$ drücken wir zunächst durch D'_{kl} aus:

$$\gamma^\sigma D_{\sigma\alpha} = \sum_{kl} e_k e_l \alpha_k h_{l\alpha} D'_{kl}. \tag{51}$$

Die hier auftretende Summe $\sum\limits_k e_k \alpha_k D'_{kl}$ läßt sich mit Hilfe von (48) berechnen, wobei die zyklische Symmetrie des Riemannschen Tensors zu beachten ist. Man bekommt

$$\sum_k e_k \alpha_k D'_{kl} = \sum_k e_k \alpha_k \left(-\frac{1}{2} R'_{kl} + \frac{2\pi i e}{hc} M'_{kl}\right), \tag{52}$$

wo

$$R'_{kl} = -\sum_i e_i \gamma_{ikil} \tag{53}$$

die Beinkomponenten des verjüngten Riemannschen Tensors bezeichnen. Setzt man (52) in (51) ein, so bekommt man

$$\gamma^\sigma D_{\sigma\alpha} = \gamma^\varrho \left(-\frac{1}{2} R_{\varrho\alpha} + \frac{2\pi i e}{hc} M_{\varrho\alpha}\right). \tag{51*}$$

Wir haben somit für die Divergenz des Tensors $A^{\sigma\cdot}_{\cdot\alpha}$ den Ausdruck

$$\frac{1}{\sqrt{g}} \frac{\partial}{\partial x^\sigma} \left(\sqrt{g}\, A^{\sigma\cdot}_{\cdot\alpha}\right) - \Gamma^\sigma_{\alpha\varrho} A^{\varrho\cdot}_{\cdot\sigma} = S^\varrho \left(-\frac{1}{2} R_{\varrho\alpha} + \frac{2\pi i e}{hc} M_{\varrho\alpha}\right). \tag{54}$$

Setzen wir

$$\frac{ch}{2\pi i} A^{\sigma\cdot}_{\cdot\alpha} = W^{\sigma\cdot}_{\cdot\alpha} = T^{\sigma\cdot}_{\cdot\alpha} + i U^{\sigma\cdot}_{\cdot\alpha}, \tag{55}$$

wo $T^{\sigma\cdot}_{\cdot\alpha}$ und $U^{\sigma\cdot}_{\cdot\alpha}$ den reellen und den imaginären Teil des komplexen Tensors $W^{\sigma\cdot}_{\cdot\alpha}$ bezeichnen, so läßt sich die Formel (54) in der Form

$$\nabla_\sigma W^{\sigma\cdot}_{\cdot\alpha} = S^\varrho \left(e M_{\varrho\alpha} - \frac{hc}{4\pi i} R_{\varrho\alpha}\right) \tag{56}$$

schreiben, oder, wenn man den reellen und den imaginären Teil trennt,

$$\left.\begin{aligned} \nabla_\sigma T^{\sigma\cdot}_{\cdot\alpha} &= e S^\varrho M_{\varrho\alpha}, \\ \nabla_\sigma U^{\sigma\cdot}_{\cdot\alpha} &= \frac{hc}{4\pi} S^\varrho R_{\varrho\alpha}. \end{aligned}\right\} \tag{57}$$

Die zweite dieser Gleichungen ist eine leicht zu beweisende Identität, denn der Tensor $U^{\sigma\cdot}_{\cdot\alpha}$ ist gleich

$$U^{\sigma\cdot}_{\cdot\alpha} = -\frac{hc}{4\pi} \nabla_\alpha S^\sigma, \tag{58}$$

und die Divergenz des Vektors S^σ verschwindet nach (29).

Die Gleichung (57) sagt aus, daß die Divergenz des Tensors $T^{\sigma\cdot}_{\cdot\alpha}$ gleich der Lorentzkraft ist. Wir können deshalb $T^{\sigma\cdot}_{\cdot\alpha}$ als Energietensor deuten. Die Gleichungen (57) sind dann die Bewegungsgleichungen der allgemeinen Relativitätstheorie. Vielleicht wäre es konsequenter, nicht den reellen Teil $T^{\sigma\cdot}_{\cdot\alpha}$, sondern den gesamten komplexen Tensor $W^{\sigma\cdot}_{\cdot\alpha}$ als Energietensor zu deuten; auf die Frage, welche dieser Deutungen vorzuziehen ist, gehen wir hier nicht ein.

Bemerkenswert ist hier das Auftreten des elektromagnetischen Tensors $M_{\varrho\alpha}$ neben dem Riemannschen Tensor $R_{\varrho\alpha}$ in der Form einer hermitischen Matrix

$$\left\| R_{\varrho\alpha} - \frac{4\pi i e}{h c} M_{\varrho\alpha} \right\|.$$

8. Um aus den gewonnenen Resultaten die quantenmechanischen Bewegungsgleichungen abzuleiten, welche denjenigen eines Massenpunktes (geodätische Linie) entsprechen, verfahren wir folgendermaßen.

Wir wählen ein im Bereich der räumlichen Variablen x_1, x_2, x_3 vollständiges System von Funktionen:

$$\psi_n(x_0\, x_1\, x_2\, x_3;\, \zeta) \qquad (\zeta = 1, 2, 3, 4), \tag{59}$$

deren jede der Diracschen Gleichung genügt* und durch die Forderung

$$\iiint \bar{\psi}_n \gamma^0 \psi_n \sqrt{g}\, d x_1\, d x_2\, d x_3 = 1 \tag{60}$$

normiert ist. Wegen (26) und (27) folgt aus dem Bestehen dieser Gleichung für einen speziellen Wert von x_0 ihre Gültigkeit für jeden anderen Wert von x_0. Das Matrixelement für einen Operator L definieren wir durch die Formel

$$L_{mn} = \iiint \bar{\psi}_m L\, \psi_n \sqrt{g}\, d x_1\, d x_2\, d x_3. \tag{61}$$

Wir beachten, daß die Ausführungen des vorigen Paragraphen und speziell die Gleichung (54) ungeändert bleiben, wenn wir in $A^{\sigma\cdot}_{\cdot\alpha}$ und in S^{ϱ} ψ durch ψ_n und $\bar{\psi}$ durch $\bar{\psi}_m$, also durch zwei verschiedene Lösungen der Diracschen Gleichung ersetzen. Wir schreiben jetzt die Gleichung (54) in der Form

$$\begin{aligned} \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^{\sigma}} \left(\bar{\psi}_m \sqrt{g}\, \gamma^{\sigma} D_{\alpha} \psi_n\right) &= \Gamma^{\sigma}_{\alpha\varrho}\, \bar{\psi}_m \gamma^{\varrho} D_{\sigma} \psi_n \\ &+ \bar{\psi}_m \gamma^{\varrho} \left(-\frac{1}{2} R_{\varrho\alpha} + \frac{2\pi i e}{h c} M_{\varrho\alpha}\right) \psi_n. \end{aligned} \tag{62}$$

* Vgl. hierzu V. Fock, ZS. f. Phys. **49**, 323, 1928 sowie **55**, 127, 1929.

Multiplizieren wir (62) mit $\sqrt{g}\,dx_1\,dx_2\,dx_3$, und integrieren wir über den ganzen Raum, so bleibt von der Summe auf der linken Seite von (62) nur ein einziges Glied übrig, und wir erhalten

$$\frac{d}{dx^0}\left\{\iiint \overline{\psi}_m \gamma^0 D_\alpha \psi_n \sqrt{g}\,dx_1\,dx_2\,dx_3\right\}$$

$$= \iiint \overline{\psi}_m \left[\Gamma^\sigma_{\alpha\varrho}\gamma^\varrho D_\sigma + \gamma^\varrho\left(-\frac{1}{2}R_{\varrho\alpha} + \frac{2\pi i e}{hc}M_{\varrho\alpha}\right)\right]\psi_n \cdot \sqrt{g}\,dx_1\,dx_2\,dx_3, \quad (63)$$

welche Gleichung wir symbolisch auch in der Form

$$\frac{d}{dx^0}(\gamma^0 D_\alpha) = \Gamma^\sigma_{\alpha\varrho}\gamma^\varrho D_\sigma + \gamma^\varrho\left(-\frac{1}{2}R_{\varrho\alpha} + \frac{2\pi i e}{hc}M_{\varrho\alpha}\right) \quad (64)$$

schreiben können, oder auch, wenn wir

$$P_\alpha = \frac{h}{2\pi i}D_\alpha \quad (65)$$

setzen, in der Form

$$\frac{d}{dx^0}(\gamma^0 P_\alpha) = \Gamma^\sigma_{\alpha\varrho}\gamma^\varrho P_\sigma + \gamma^\varrho\left(\frac{e}{c}M_{\varrho\alpha} - \frac{h}{4\pi i}R_{\varrho\alpha}\right). \quad (66)$$

Nun können wir die Operatoren γ^ϱ als Repräsentanten der klassischen Geschwindigkeit $\frac{dx^\varrho}{dx^0}$ und P_σ als diejenigen der kovarianten Bewegungsgrößen $mg_{\sigma\varrho}\frac{dx^\varrho}{ds}$ deuten. Diese Deutung ermöglicht es, den Übergang zur klassischen Theorie zu vollziehen. Tun wir dies, und vernachlässigen wir konsequenterweise das mit h behaftete Glied auf der rechten Seite, so erhalten wir genau die klassische Bewegungsgleichung für einen geladenen Massenpunkt im Gravitationsfeld und speziell — wenn kein elektromagnetisches Feld vorhanden ist — die Differentialgleichung der geodätischen Linie.

9. Der rein kovariante Tensor

$$W_{\sigma\alpha} = g_{\sigma\varrho}W^\varrho_{\cdot\alpha} = c\,\overline{\psi}\gamma_\sigma P_\alpha \psi \quad (55^*)$$

ist in bezug auf seine Indizes nicht symmetrisch. Wegen der Bedeutung der Operatoren $c\gamma_\sigma$ und P_α (Geschwindigkeit und Bewegungsgröße) korrespondiert die quantenmechanische Größe $W_{\sigma\alpha}$ der klassischen Größe $\varrho_0 u_\sigma u_\alpha$

$$W_{\sigma\alpha} \rightarrow \varrho_0 u_\sigma u_\alpha, \quad (67)$$

wo u_σ die klassische kovariante Komponente der Vierergeschwindigkeit und ϱ_0 die Ruhdichte der Materie bezeichnet. Die Größe $\varrho_0 u_\sigma u_\alpha$ ist aber in bezug auf die Indizes symmetrisch.

Die Diracsche Gleichung (27) läßt sich aus einem Variationsprinzip ableiten, welches mit Hilfe des Energietensors wie folgt formuliert werden kann:

$$\delta \iiiint (W^\sigma_{\cdot\sigma} - mc^2\overline{\psi}\alpha_4\psi)\sqrt{g}\,dx_0\,dx_1\,dx_2\,dx_3 = 0. \quad (68)$$

Diese Gleichung liefert eine einfache physikalische Deutung der Invariante $m\overline{\psi}\alpha_4\psi$ als Ruhdichte der Materie.

10. Wir wollen nun die Ergebnisse unserer Untersuchung zusammenfassen.

Als Ausgangspunkt diente uns der Begriff der Parallelverschiebung eines Halbvektors. Mit Hilfe dieses Begriffs konnte das Auftreten der Potentiale φ_α neben den Impulsen p_α in der Diracschen Gleichung rein geometrisch gedeutet werden; die rein formelle Übertragung des Ausdrucks $p_\alpha - \frac{e}{c}\varphi_\alpha$ aus der klassischen Mechanik in die Quantenmechanik wurde damit überflüssig. Ferner erlaubte uns der genannte Begriff eine ungezwungene Einordnung der Potentiale in das geometrische Schema der allgemeinen Relativitätstheorie, was für die Aufstellung einer einheitlichen Theorie der Elektrizität und der Gravitation von Nutzen sein kann.

Ferner wurden die Diracschen Gleichungen in der allgemeinen Relativitätstheorie aufgestellt, die in bezug auf die Wahl der Koordinaten und der „Beine" invariant sind. Als Nebenresultat ergab sich daraus eine explizite Darstellung des Diracschen Operators in krummlinigen orthogonalen Koordinaten. Es wurde ein Tensor konstruiert, dessen Divergenz gleich der Lorentzkraft ist; dieser Tensor wurde als Energietensor und die von ihm befriedigte Gleichung als makroskopische Bewegungsgleichung gedeutet. Ferner wurden die quantenmechanischen Bewegungsgleichungen für das Elektron abgeleitet, die den klassischen Gleichungen für einen geladenen Massenpunkt oder — bei Abwesenheit eines elektromagnetischen Feldes — denjenigen einer geodätischen Linie entsprechen. Zum Schluß wurde das Variationsprinzip, aus welchem die Diracsche Gleichung abgeleitet werden kann, hingeschrieben.

Das Ziel, welchem wir nachstreben, war die Geometrisierung der Diracschen Theorie des Elektrons und ihre Einordnung in die allgemeine Relativitätstheorie. Dabei wurden die Schwierigkeiten, die der Diracschen Theorie anhaften — wie das Auftreten der negativen Energiewerte und eine nichtverschwindende Wahrscheinlichkeit einer Umladung des Elektrons —, überhaupt nicht berührt. Vielleicht dürften aber unsere Betrachtungen auf indirektem Wege zur Lösung dieser Schwierigkeiten beitragen, indem sie zeigen, was die ursprüngliche, unveränderte Diracsche Theorie leisten kann.

Leningrad, Physikalisches Institut der Universität, Mai/Juni 1929.

Nachtrag.

Nach dem Abschluß dieser Arbeit habe ich die sehr interessante Arbeit von H. Weyl * kennengelernt. Weyls mathematische Grundidee ist im wesentlichen mit dem Begriff der Parallelübertragung eines Halbvektors identisch. Der physikalische Inhalt von Weyls Arbeit ist jedoch von demjenigen meiner Arbeit völlig verschieden.

Die wesentlichen Züge der Auffassung von Weyl können wie folgt zusammengefaßt werden.

1. Weyl betrachtet die Diracsche Gleichung als eine Wellengleichung nicht für das Elektron, sondern für das System Elektron-Proton.

2. In den Gravitationszusatzgliedern glaubt Weyl einen Ersatz für das Glied $mc\alpha_4$ zu finden, welch letzteres einfach gestrichen wird.

Beide Thesen können, meiner Ansicht nach, kaum aufrechterhalten werden, denn sie stoßen auf wesentliche Schwierigkeiten, auf die ich hier aufmerksam machen möchte.

Die aus der Diracschen Gleichung folgenden quantenmechanischen Bewegungsgleichungen sind ein vollständiges Analogon zu den klassischen Bewegungsgleichungen für einen geladenen Massenpunkt (und nicht etwa für ein System aus zwei Körpern), wie bereits in meiner früheren Arbeit** gezeigt wurde.

Die Diracsche Gleichung, und zwar mit dem Gliede $mc\alpha_4$, eignet sich vollkommen zur Beschreibung einer kräftefreien Bewegung eines Elektrons als Welle im Sinne der ursprünglichen de Broglieschen Auffassung.

Die von Weyl vorgenommene Zerlegung des Stromvektors S in zwei Summanden $S^{(+)}$ und $S^{(-)}$, die als Strom positiver und negativer Elektrizität gedeutet werden, kann nicht aufrechterhalten werden, denn diese Summanden sind Nullvektoren, und nur ihre Summe $S = S^{(+)} + S^{(-)}$ ist ein zeitartiger Vektor***. Der Strom ist aber eine statistisch-

* H. Weyl, Gravitation and the Electron, Proc. Nat. Acad. Amer. **15**, 323, 1929.

** V. Fock, Über den Begriff der Geschwindigkeit usw., ZS. f. Phys. **55**, 127, 1929.

*** Beweis: Der zeitartige Charakter von S folgt aus der Identität (8) (wo jetzt S_i statt A_i zu lesen ist), denn sie ergibt

$$S_0^2 - S_1^2 - S_2^2 - S_3^2 = S_4^2 + S_5^2, \tag{*}$$

$S_i^{(+)}$ bzw. $S_i^{(-)}$ erhält man aus S_i, wenn man ψ_3 und ψ_4 bzw. ψ_1 und ψ_2 gleich Null setzt; in beiden Fällen verschwindet S_4 und S_5, also auch die linke Seite von (*), was zu beweisen war.

makroskopische Größe und muß als solche denselben Charakter haben wie in der klassischen Theorie, also notwendig zeitartig sein.

Die Gleichungen von Weyl sollen das System Elektron-Proton beschreiben; man darf daher verlangen, daß sie die Energieniveaus des Wasserstoffatoms richtig wiedergeben. Wegen der Weglassung des Gliedes $mc\alpha_4$ ist das aber kaum möglich und jedenfalls nicht bewiesen.

Die von Weyl als Ersatz der Masse gedeuteten Gravitationsglieder [„Beinvektor“ f_i in unserer Formel (35)] können zum Verschwinden gebracht werden, sobald ein System von Normalenkongruenzen existiert und speziell im Fall einer sphärischen Symmetrie, sowie im statischen Falle einer axialen Symmetrie. Von dem System Elektron-Proton dürfte man aber einen hohen Grad von Symmetrie erwarten.

Endlich bleibt es ganz unklar, wie eigentlich die Konstanten m und M — die Massen des Elektrons und des Protons — aus den Gravitationsgliedern zum Vorschein kommen sollen.

Wegen dieser Schwierigkeiten kann ich Weyls Versuch, das quantenmechanische Problem der Masse sowie das Zweikörperproblem anzugreifen, nicht als gelungen betrachten. Der allgemeinen Idee von Weyl, daß beide Probleme miteinander und mit der Gravitation eng verbunden sind, stimme ich dagegen gerne bei.

Zum Schluß möchte ich einige allgemeine Bemerkungen über den physikalischen Inhalt der Diracschen Gleichungen und über das quantenmechanische Zweikörperproblem machen.

Nach meiner Auffassung wird durch die Diracsche Gleichung nur das Elektron quantenmechanisch, die übrige Welt dagegen (vielleicht auch die Masse des Elektrons) makroskopisch beschrieben. Zur übrigen Welt wird hier auch das Proton gerechnet. Die Lösung des Zweikörperproblems muß darin bestehen, daß man eine quantenmechanische Beschreibung des Elektrons, des Protons, des elektromagnetischen Feldes und der Masse findet. Das quantenmechanische Problem der Masse scheint mir unangreifbar, solange man nur einen Körper betrachtet. Für die makroskopische Beschreibung der Gravitation und der Elektrizität scheint dagegen das quantenmechanische Einkörperproblem gute Dienste leisten zu können.

Chapter 12
Hermann Weyl (1929): Elektron und Gravitation

Hermann Weyl (1929). Elektron und Gravitation, *Zeitschrift für Physik*, 56: 330–352.

Elektron und Gravitation. I.

Von **Hermann Weyl** in Princeton, N. J.

(Eingegangen am 8. Mai 1929).

Einleitung. Verhältnis der allgemeinen Relativitätstheorie zu den quantentheoretischen Feldgleichungen des spinnenden Elektrons: Masse, Eichinvarianz, Fernparallelismus. Zu erwartende Modifikationen der Diracschen Theorie. — I. Zweikomponententheorie: Die Wellenfunktion ψ hat nur zwei Komponenten. — § 1. Bindung der Transformation der ψ an die Lorentztransformation des normalen Achsenkreuzes in der vierdimensionalen Welt. Asymmetrie von Zukunft und Vergangenheit, von rechts und links. — § 2. In der allgemeinen Relativitätstheorie wird die Metrik in einem Weltpunkt festgelegt durch ein normales Achsenkreuz. Komponenten von Vektoren relativ zu den Achsen und den Koordinaten. Kovariante Differentiation von ψ. — § 3. Allgemein invariante Fassung der Diracschen Wirkungsgröße, welche für das Wellenfeld der Materie charakteristisch ist. — § 4. Die differentiellen Erhaltungssätze von Energie und Impuls und die Symmetrie des Impulstensors folgen aus der doppelten Invarianz: 1. gegenüber Koordinatentransformation, 2. gegenüber Drehungen des Achsenkreuzes. Impuls und Impulsmoment der Materie. — § 5. Einsteins klassische Gravitationstheorie in der neuen analytischen Formulierung. Gravitationsenergie. — § 6. Das elektromagnetische Feld. Aus der Unbestimmtheit des Eichfaktors in ψ ergibt sich die Notwendigkeit der Einführung der elektromagnetischen Potentiale. Eichinvarianz und Erhaltung der Elektrizität. Das Raumintegral der Ladung. Einführung der Masse. Diskussion und Zurückweisung einer anderen Möglichkeit, in welcher die Elektrizität nicht als Begleitphänomen der Materie, sondern der Gravitation erscheint.

Einleitung.

In dieser Arbeit entwickle ich in ausgeführter Form eine Gravitation, Elektrizität und Materie umfassende Theorie, von der eine kurze Skizze in den Proc. Nat. Acad., April 1929, erschienen ist. Es ist von verschiedenen Autoren der Zusammenhang der Einsteinschen Theorie des Fernparallelismus mit der Spintheorie des Elektrons bemerkt worden *. Trotz gewisser formaler Übereinstimmungen unterscheidet sich mein Ansatz in radikaler Weise dadurch, daß ich den Fernparallelismus ablehne und an Einsteins klassischer Relativitätstheorie der Gravitation festhalte.

Um zweier Gründe willen verspricht die Adaption der Pauli-Diracschen Theorie des spinnenden Elektrons an die allgemeine Relativität zu physikalisch fruchtbaren Ergebnissen zu führen. 1. Die Diracsche Theorie, in welcher das Wellenfeld des Elektrons durch ein Potential ψ mit vier Komponenten beschrieben wird, gibt doppelt zu viel Energieniveaus; man sollte darum, ohne die relativistische Invarianz preiszugeben, zu den zwei Komponenten der Paulischen Theorie zurück-

* E. Wigner, ZS. f. Phys. **53**, 592, 1929; u. a.

kehren können. Daran hindert das die Masse m des Elektrons als Faktor enthaltende Glied der Diracschen Wirkungsgröße. Masse ist aber ein Gravitationseffekt; es besteht so die Hoffnung, für dieses Glied in der Gravitationstheorie einen Ersatz zu finden, der die gewünschte Korrektur herbeiführt. 2. Die Diracschen Feldgleichungen für ψ zusammen mit den Maxwellschen Gleichungen für die vier Potentiale f_p des elektromagnetischen Feldes haben eine Invarianzeigenschaft, die in formaler Hinsicht derjenigen gleicht, die ich in meiner Theorie von Gravitation und Elektrizität vom Jahre 1918 als Eichinvarianz bezeichnet hatte; die Gleichungen bleiben ungeändert, wenn man gleichzeitig

$$\psi \text{ durch } e^{i\lambda}\cdot\psi \quad \text{und} \quad f_p \text{ durch } f_p - \frac{\partial\lambda}{\partial x_p}$$

ersetzt, unter λ eine willkürliche Ortsfunktion in der vierdimensionalen Welt verstanden. Dabei ist in f_p der Faktor $\frac{e}{ch}$ aufgenommen ($-e$ Ladung des Elektrons, c Lichtgeschwindigkeit, $\frac{h}{2\pi}$ Wirkungsquantum). Auch die Beziehung dieser „Eichinvarianz" zum Erhaltungssatz der Elektrizität bleibt unangetastet. Es ist aber ein wesentlicher und für den Anschluß an die Erfahrung bedeutungsvoller Unterschied, daß der Exponent des Faktors, den ψ annimmt, nicht reell, sondern rein imaginär ist. ψ übernimmt jetzt die Rolle, welche in jener alten Theorie das Einsteinsche ds spielte. Es scheint mir darum dieses nicht aus der Spekulation, sondern aus der Erfahrung stammende neue Prinzip der Eichinvarianz zwingend darauf hinzuweisen, daß das elektrische Feld ein notwendiges Begleitphänomen nicht des Gravitationsfeldes, sondern des materiellen, durch ψ dargestellten Wellenfeldes ist. Da die Eichinvarianz eine willkürliche Funktion λ einschließt, hat sie den Charakter „allgemeiner" Relativität und kann natürlich nur in ihrem Rahmen verstanden werden.

An den Fernparallelismus vermag ich aus mehreren Gründen nicht zu glauben. Erstens sträubt sich mein mathematisches Gefühl a priori dagegen, eine so künstliche Geometrie zu akzeptieren; es fällt mir schwer, die Macht zu begreifen, welche die lokalen Achsenkreuze in den verschiedenen Weltpunkten in ihrer verdrehten Lage zu starrer Gebundenheit aneinander hat einfrieren lassen. Es kommen, wie ich glaube, zwei gewichtige physikalische Gründe hinzu. Gerade dadurch, daß man den Zusammenhang zwischen den lokalen Achsenkreuzen löst, verwandelt sich der Eichfaktor $e^{i\lambda}$, der in der Größe ψ willkürlich bleibt, notwendig aus einer Konstante in eine willkürliche Ortsfunktion; d. h. nur durch diese Lockerung wird die tatsächlich bestehende Eichinvarianz verständlich. Und zweitens ist die Möglichkeit, die Achsenkreuze an verchiedenen Stellen unabhängig voneinander zu drehen, wie wir im folgenden sehen werden, gleichbedeutend mit der Symmetrie des Energieimpulstensors oder mit der Gültigkeit des Erhaltungssatzes für das Impulsmoment.

Bei jedem Versuch zur Aufstellung der quantentheoretischen Feldgleichungen muß man im Auge haben, daß diese nicht direkt mit der Erfahrung verglichen werden können, sondern erst nach ihrer Quantisierung die Unterlage liefern für die statistischen Aussagen über das Verhalten der materiellen Teilchen und Lichtquanten. Die Dirac-Maxwellsche Theorie in ihrer bisherigen Form enthält nur die elektromagnetischen Potentiale f_p und das Wellenfeld ψ des Elektrons. Zweifellos muß das Wellenfeld ψ' des Protons hinzugefügt werden. Und zwar werden in den Feldgleichungen ψ, ψ' und f_p Funktionen derselben vier Raum-Zeitkoordinaten sein, man wird vor der Quantisierung nicht etwa verlangen dürfen, daß ψ Funktion eines Weltpunktes (t, xyz) und ψ' Funktion eines davon unabhängigen Weltpunktes $(t', x'y'z')$ ist. Es ist naheliegend, zu erwarten, daß von den beiden Komponentenpaaren der Diracschen Größe das eine dem Elektron, das andere dem Proton zugehört. Ferner werden zwei Erhaltungssätze der Elektrizität auftreten müssen, die (nach der Quantisierung) besagen, daß die Anzahl der Elektronen wie der Protonen konstant bleibt. Ihnen wird eine zweifache, zwei willkürliche Funktionen involvierende Eichinvarianz entsprechen müssen.

Wir prüfen zunächst die Sachlage in der speziellen Relativitätstheorie daraufhin, ob und inwieweit bereits die formalen Erfordernisse der Gruppentheorie, noch ganz abgesehen von den mit der Erfahrung in Einklang zu bringenden dynamischen Differentialgleichungen, die Erhöhung der Komponentenzahl ψ von zwei auf vier notwendig machen. Wir werden sehen, daß man mit zwei Komponenten auskommt, wenn die Symmetrie von links und rechts aufgehoben wird.

Zweikomponententheorie.

§ 1. Transformationsgesetz von ψ. Führt man im Raume mit den kartesischen Koordinaten x, y, z homogene projektive Koordinaten x_α ein:

$$x = \frac{x_1}{x_0}, \quad y = \frac{x_2}{x_0}, \quad z = \frac{x_3}{x_0},$$

so lautet die Gleichung der Einheitskugel

$$-x_0^2 + x_1^2 + x_2^2 + x_3^2 = 0. \tag{1}$$

Projiziert man sie vom Südpol auf die Äquatorebene $z = 0$, die als Träger der komplexen Variablen

$$x + iy = \zeta = \frac{\psi_2}{\psi_1}$$

betrachtet wird, so gelten die Gleichungen

$$\left.\begin{array}{ll} x_0 = \overline{\psi}_1 \psi_1 + \overline{\psi}_2 \psi_2, & x_1 = \overline{\psi}_1 \psi_2 + \overline{\psi}_2 \psi_1, \\ x_2 = i(-\overline{\psi}_1 \psi_2 + \overline{\psi}_2 \psi_1), & x_3 = \overline{\psi}_1 \psi_1 - \overline{\psi}_2 \psi_2. \end{array}\right\} \tag{2}$$

x_α sind Hermitesche Formen von ψ_1, ψ_2. Die Variablen ψ_1, ψ_2 sowie die Koordinaten x_α kommen hier nur ihrem Verhältnis nach in Frage. Eine homogene lineare Transformation von ψ_1, ψ_2 (mit komplexen Koeffizienten) bewirkt eine lineare, reelle Transformation unter den Koordinaten x_α: sie stellt eine Kollineation dar, welche die Einheitskugel in sich überführt und auf der Einheitskugel den Drehsinn ungeändert läßt. Es ist leicht zu zeigen und wohl bekannt, daß man auf diese Weise jede derartige Kollineation einmal und nur einmal erhält.

Vom homogenen Standpunkt zum inhomogenen übergehend, fasse man jetzt x_α als Koordinaten in der vierdimensionalen Welt und (1) als die Gleichung des „Lichtkegels" auf; und man beschränke sich auf solche lineare Transformationen U von ψ_1, ψ_2, deren Determinante den absoluten Betrag 1 hat. U bewirkt an den x_α eine Lorentztransformation, d. i. eine reelle homogene lineare Transformation, welche die Form

$$-x_0^2 + x_1^2 + x_2^2 + x_3^2$$

in sich überführt. Doch lehren die Formel für x_0 und unsere Bemerkung über die Erhaltung des Drehungssinnes auf der Kugel ohne weiteres, daß wir unter den Lorentztransformationen nur die ein einziges in sich abgeschlossenes Kontinuum bildenden Λ bekommen, welche 1. Vergangenheit und Zukunft nicht vertauschen und 2. die Determinante $+1$, nicht -1, besitzen; diese freilich ohne Ausnahme. Durch Λ ist die lineare Transformation U der ψ nicht eindeutig festgelegt, sondern es bleibt ein willkürlicher konstanter Faktor $e^{i\lambda}$ vom absoluten Betrage 1 zur Disposition. Man kann ihn normalisieren durch die Forderung, daß die Determinante von U gleich 1 sei, aber selbst dann bleibt eine Doppeldeutigkeit zurück. An der Einschränkung 1. möchte man festhalten; es ist eine der hoffnungsvollsten Seiten der ψ-Theorie, daß sie der Wesensverschiedenheit von Vergangenheit und Zukunft Rechnung tragen kann. Die Einschränkung 2. hebt die Gleichberechtigung von links und rechts auf.

Nur diese tatsächlich in der Natur bestehende Symmetrie von rechts und links wird uns zwingen (Teil II), ein zweites Paar von ψ-Komponenten einzuführen.

Die Hermitesche Konjugierte einer Matrix $A = \|a_{ik}\|$ werde mit A^* bezeichnet:

$$a^*_{ik} = \bar{a}_{ki}.$$

S_α sei die Koeffizientenmatrix der Hermiteschen Form der Variablen ψ_1, ψ_2, durch welche in (2) die Koordinate x_α dargestellt wird:

$$x_\alpha = \psi^* S_\alpha \psi; \tag{3}$$

hier bedeutet ψ die Spalte ψ_1, ψ_2. S_0 ist die Einheitsmatrix; es gelten die Gleichungen

$$S_1^2 = 1, \quad S_2 S_3 = i S_1 \tag{4}$$

und die daraus durch zyklische Vertauschung der Indizes 1, 2, 3 hervorgehenden.

Es ist formal etwas bequemer, die reelle Zeitkoordinate x_0 durch die imaginäre $i x_0$ zu ersetzen. Die Lorentztransformationen erscheinen dann als orthogonale Transformationen der vier Größen

$$x(0) = i x_0, \quad x(\alpha) = x_\alpha \qquad [\alpha = 1, 2, 3].$$

Statt (3) schreibe man

$$x(\alpha) = \psi^* S(\alpha)\, \psi. \tag{5}$$

Das Transformationsgesetz der ψ-Komponenten besteht darin, daß sie unter dem Einfluß einer Transformation Λ der Weltkoordinaten $x(\alpha)$ sich so umsetzen, daß die Größen (5) die Transformation Λ erleiden. Eine Größe von dieser Art stellt, wie sich aus dem Spinphänomen ergeben hat, das Wellenfeld eines materiellen Teilchens dar. $x(\alpha)$ sind die Koordinaten in einem „normalen Achsenkreuz" $\mathbf{e}(\alpha)$; $\mathbf{e}(1)$, $\mathbf{e}(2)$, $\mathbf{e}(3)$ sind reelle raumartige Vektoren, welche ein kartesisches Linkskoordinatensystem bilden, $\dfrac{\mathbf{e}(0)}{i}$ ist ein reeller zeitartiger, in die Zukunft gerichteter Weltvektor. Die Transformation Λ beschreibt den Übergang von einem solchen normalen Achsenkreuz zu einem anderen gleichberechtigten, der weiterhin kurz als Drehung des Achsenkreuzes bezeichnet werden möge. Wir bekommen dieselben Koeffizienten $c(\alpha\beta)$, ob wir die Transformation Λ an den Grundvektoren des Achsenkreuzes oder den Koordinaten ausdrücken:

$$\mathbf{x} = \sum_\alpha x(\alpha)\, \mathbf{e}(\alpha) = \sum_\alpha x'(\alpha)\, \mathbf{e}'(\alpha),$$

$$\mathbf{e}'(\alpha) = \sum_\beta c(\alpha\beta)\, \mathbf{e}(\beta), \quad x'(\alpha) = \sum_\beta c(\alpha\beta)\, x(\beta);$$

das folgt aus dem orthogonalen Charakter von Λ.

Für das Folgende ist es nötig, die infinitesimale Transformation

$$d\psi = dE \,.\, \psi \tag{6}$$

zu berechnen, welche einer beliebigen infinitesimalen Drehung $d\Omega$:

$$dx(\alpha) = \sum_{\beta} do(\alpha\beta) \,.\, x(\beta),$$

entspricht. Die $do(\alpha\beta)$ bilden eine schiefsymmetrische Matrix. Die Transformation (6) ist so normiert gedacht, daß die Spur von dE gleich 0 wird. Die Matrix dE hängt linear homogen von den $do(\alpha\beta)$ ab; wir schreiben daher

$$dE = \tfrac{1}{2} \sum_{\alpha\beta} do(\alpha\beta) \,.\, A(\alpha\beta) = \sum do(\alpha\beta) \,.\, A(\alpha\beta).$$

Die letzte Summe soll nur über die Paare

$$(\alpha\beta) = (0\,1),\ (0\,2),\ (0\,3);\quad (2\,3),\ (3\,1),\ (1\,2)$$

erstreckt werden. $A(\alpha\beta)$ hängt natürlich schiefsymmetrisch von α und β ab. Es darf nicht vergessen werden, daß die Koeffizienten $do(\alpha\beta)$ für die ersten drei Paare $(\alpha\beta)$ rein imaginär, für die letzten drei Paare reell, im übrigen aber willkürlich sind. Man findet

$$A(2\,3) = -\frac{1}{2i} S(1), \quad A(0\,1) = \frac{1}{2i} S(1) \tag{7}$$

und zwei analoge Paare von Gleichungen, die daraus durch zyklische Vertauschung der Indizes 1, 2, 3 entstehen. Zur Bestätigung hat man lediglich auszurechnen, daß die infinitesimalen Transformationen dE

$$d\psi = \frac{1}{2i} S(1)\,\psi \quad \text{und} \quad d\psi = \frac{1}{2} S(1)\,\psi$$

die infinitesimalen Drehungen

$$dx(0) = 0, \quad dx(1) = 0, \quad dx(2) = -x(3), \quad dx(3) = x(2)$$

bzw.

$$dx(0) = ix(1), \quad dx(1) = -ix(0), \quad dx(2) = 0, \quad dx(3) = 0$$

hervorbringen.

§ 2. Metrik und Parallelverschiebung. Wir gehen über zur allgemeinen Relativitätstheorie. Die Metrik in einem Weltpunkte P beschreiben wir durch Angabe eines lokalen normalen Achsenkreuzes $\mathfrak{e}(\alpha)$. Nur die Klasse der normalen Achsenkreuze — welche durch die Gruppe der Drehungen Λ miteinander verbunden sind — ist durch die Metrik bestimmt; durch einen Akt der Willkür wird ein einzelnes Individuum aus dieser Klasse ausgesucht. Die Gesetze sind demnach invariant gegenüber beliebigen Drehungen der lokalen Achsenkreuze; dabei ist die Drehung des Achsenkreuzes in dem von P verschiedenen Punkt P' unabhängig von der Drehung in P. $\psi_1(P)$, $\psi_2(P)$ seien die Komponenten des Materiepotentials im Punkte P relativ zum daselbst gewählten lokalen Achsenkreuz $\mathbf{e}(\alpha)$. Ein Vektor $\boldsymbol{t}$ in P kann in der Form geschrieben werden

$$\boldsymbol{t} = \sum_\alpha t(\alpha)\,\mathbf{e}(\alpha);$$

die Zahlen $t(\alpha)$ sind seine Komponenten im Achsenkreuz.

Zur analytischen Darstellung bedürfen wir ferner eines Koordinatensystems x_p; x_p sind irgend vier stetige Ortsfunktionen in der Welt, deren Werte die verschiedenen Weltpunkte voneinander zu unterscheiden gestatten. Die Gesetze sind demnach invariant gegenüber beliebigen Koordinatentransformationen. $e^p(\alpha)$ mögen die Komponenten von $\mathbf{e}(\alpha)$ im Koordinatensystem sein. Diese $4 \cdot 4$ Größen $e^p(\alpha)$ beschreiben das Gravitationsfeld. Die kontravarianten Komponenten t^p eines Vektors $\boldsymbol{t}$ im Koordinatensystem hängen mit seinen Komponenten $t(\alpha)$ im Achsenkreuz durch die Gleichungen zusammen:

$$t^p = \sum_\alpha t(\alpha) \cdot e^p(\alpha).$$

Andererseits berechnen sich die $t(\alpha)$ aus seinen kovarianten Komponenten t_p im Koordinatensystem vermöge

$$t(\alpha) = \sum_p t_p \cdot e^p(\alpha).$$

Diese Gleichungen regeln die Verwandlung der Indizes. Die auf das Achsenkreuz bezüglichen griechischen Indizes habe ich als Argumente geschrieben, weil hier zwischen Hoch- und Tiefstellung nicht zu unterscheiden ist. Die Verwandlung im umgekehrten Sinne geschieht mittels der zu $\| e^p(\alpha) \|$ inversen Matrix $\| e_p(\alpha) \|$:

$$\sum_\alpha e_p(\alpha)\, e^q(\alpha) = \delta_p^q \quad \text{und} \quad \sum_p e_p(\alpha)\, e^p(\beta) = \delta(\alpha, \beta).$$

δ ist 0 oder 1, je nachdem die Indizes übereinstimmen oder nicht. Die Regel über das Fortlassen der Summenzeichen wird fortan sowohl für die lateinischen wie die griechischen Indizes befolgt. ε sei der absolute Betrag der Determinante $|e^p(\alpha)|$. Die Division einer lateinisch benannten Größe durch ε wird, wie üblich, durch die Verwandlung des lateinischen in den entsprechenden deutschen Buchstaben bezeichnet; z. B.

$$\mathfrak{e}^p(\alpha) = \frac{e^p(\alpha)}{\varepsilon}.$$

Einen Vektor und einen Tensor kann man durch die auf das Koordinatensystem oder durch die auf das Achsenkreuz bezüglichen Komponenten beschreiben. In bezug auf die Größe ψ kann aber nur von Komponenten im Achsenkreuz die Rede sein. Denn das Transformationsgesetz ihrer Komponenten ist durch eine Darstellung der Drehungsgruppe geregelt, welche sich nicht auf die Gruppe aller linearen Transformationen ausdehnen läßt. Daher die Notwendigkeit, in der Theorie der Materie das Gravitationsfeld auf die hier geschilderte Art, statt durch die metrische Grundform

$$\sum_{p,\,q} g_{pq}\, dx_p\, dx_q,$$

analytisch darzustellen *. Übrigens ist

$$g_{pq} = e_p(\alpha)\, e_q(\alpha).$$

Die Gravitationstheorie muß nun in diese neue analytische Form umgegossen werden. Ich beginne mit den Formeln für die durch die Metrik bestimmte i n f i n i t e s i m a l e P a r a l l e l v e r s c h i e b u n g. Der Vektor $\mathbf{e}(\alpha)$ im Punkte P gehe durch sie in den Vektor $\mathbf{e}'(\alpha)$ im unendlich benachbarten Punkte P' über. Die $\mathbf{e}'(\alpha)$ bilden in P' ein normales Achsenkreuz, das aus dem lokalen Achsenkreuz $\mathbf{e}(\alpha) = \mathbf{e}(\alpha; P')$ daselbst durch eine infinitesimale Drehung $d\Omega$ hervorgeht:

$$\delta\mathbf{e}(\beta) = \sum_{\gamma} do(\beta\gamma)\cdot\mathbf{e}(\gamma), \quad \delta\mathbf{e}(\beta) = \mathbf{e}'(\beta) - \mathbf{e}(\beta; P'). \qquad (8)$$

$d\Omega$ hängt linear von der Verschiebung PP' oder ihren Komponenten

$$dx_p = (dx)^p = v^p = e^p(\alpha)\, v(\alpha)$$

ab. Wir schreiben darum

$$d\Omega = \Omega_p(dx)^p, \quad do(\beta\gamma) = o_p(\beta\gamma)(dx)^p = o(\alpha;\beta\gamma)\, v(\alpha). \qquad (9)$$

Die Parallelverschiebung des Vektors $\mathbf{t}$ mit den Komponenten t^p wird, wie man weiß, durch eine Gleichung beschrieben:

$$d\mathbf{t} = -d\Gamma.\mathbf{t}, \quad \text{d. i.} \quad dt^p = -d\Gamma^p_r.t^r, \quad d\Gamma^p_r = \Gamma^p_{rq}(dx)^q,$$

in welcher die sowohl von $\mathbf{t}$ wie von der Verschiebung dx unabhängigen Größen Γ^p_{rq} symmetrisch in r und q sind. Wir haben also

$$\mathbf{e}'(\beta) - \mathbf{e}(\beta) = -d\Gamma.\mathbf{e}(\beta).$$

Daneben gilt die Gleichung (8). Subtraktion der beiden Differenzen auf der linken Seite ergibt das Differential $d\mathbf{e}(\beta) = \mathbf{e}(\beta; P') - \mathbf{e}(\beta; P)$:

$$d\,\mathbf{e}^p(\beta) + d\Gamma^p_r e^r(\beta) = -do(\beta\gamma).e^p(\gamma),$$

$$\frac{\partial e^p(\beta)}{\partial x_q}\cdot e^q(\alpha) + \Gamma^p_{rq}\, e^r(\beta)\, e^q(\alpha) = -o(\alpha;\beta\gamma)\, e^p(\gamma).$$

---

* In formaler Übereinstimmung mit E i n s t e i n s neueren Arbeiten über Gravitation und Elektrizität, Sitzungsber. Preuß. Ak. Wissensch. 1928, S. 217, 224; 1929, S. 2. E i n s t e i n gebraucht den Buchstaben h statt e.

Man kann hier die o eliminieren und erhält die bekannten Gleichungen zur Bestimmung von Γ, wenn man ausdrückt, daß $o(\alpha; \beta\gamma)$ schiefsymmetrisch ist in bezug auf β und γ. Man eliminiert die Γ und berechnet o, indem man Gebrauch davon macht, daß Γ^p_{rq} symmetrisch in bezug auf r und q oder

$$\Gamma^p(\beta, \alpha) = \Gamma^p_{rq} e^r(\beta) e^q(\alpha)$$

symmetrisch in α und β ist:

$$\frac{\partial e^p(\alpha)}{\partial x_q} \cdot e^q(\beta) - \frac{\partial e^p(\beta)}{\partial x_q} \cdot e^q(\alpha) = \{o(\alpha; \beta\gamma) - o(\beta; \alpha\gamma)\} e^p(\gamma). \quad (10)$$

Die linke Seite besteht aus den Komponenten jenes gegenüber Koordinatentransformation invarianten „Kommutatorprodukts" der beiden Vektorfelder $\mathbf{e}(\alpha)$, $\mathbf{e}(\beta)$, welches die entscheidende Rolle in der Lieschen Theorie der infinitesimalen Transformationen spielt; es soll mit $[\mathbf{e}(\alpha), \mathbf{e}(\beta)]$ bezeichnet werden. Weil $o(\beta; \alpha\gamma)$ schiefsymmetrisch ist in α und γ, hat man daher

$$[\mathbf{e}(\alpha), \mathbf{e}(\beta)]^p = \{o(\alpha; \beta\gamma) + o(\beta; \gamma\alpha)\} e^p(\gamma)$$

oder

$$o(\alpha; \beta\gamma) + o(\beta; \gamma\alpha) = [\mathbf{e}(\alpha), \mathbf{e}(\beta)](\gamma). \quad (11)$$

Nimmt man in dieser Gleichung die drei zyklischen Vertauschungen von $\alpha\beta\gamma$ vor und addiert die entstehenden Gleichungen mit den Vorzeichen $+ - +$, so erhält man

$$2\,o(\alpha; \beta\gamma) = [\mathbf{e}(\alpha), \mathbf{e}(\beta)](\gamma) - [\mathbf{e}(\beta), \mathbf{e}(\gamma)](\alpha) + [\mathbf{e}(\gamma), \mathbf{e}(\alpha)](\beta).$$

$o(\alpha; \beta\gamma)$ ist also in der Tat eindeutig bestimmt. Der gefundene Ausdruck genügt allen Bedingungen, weil er, wie ohne weiteres ersichtlich, schiefsymmetrisch in β und γ ist.

Für das Folgende benötigen wir insbesondere die Verkürzung

$$o(\varrho, \varrho\alpha) = [\mathbf{e}(\alpha), \mathbf{e}(\varrho)](\varrho) = \frac{\partial e^p(\alpha)}{\partial x_p} - \frac{\partial e^p(\varrho)}{\partial x_q} \cdot e^q(\alpha) e_p(\varrho).$$

Da

$$-\varepsilon \cdot \delta\left(\frac{1}{\varepsilon}\right) = \frac{\delta \varepsilon}{\varepsilon} = e_q(\varrho) \cdot \delta e^q(\varrho)$$

ist, kommt

$$o(\varrho, \varrho\alpha) = \varepsilon \cdot \frac{\partial \mathfrak{e}^p(\alpha)}{\partial x^p}. \quad (12)$$

§ 3. Wirkung der Materie. Mit Hilfe der Parallelverschiebung kann nicht nur die kovariante Ableitung eines Vektor- oder Tensorfeldes, sondern auch diejenige des ψ-Feldes berechnet werden. $\psi_a(P)$, $\psi_a(P')$ $[a = 1, 2]$ seien die Komponenten relativ zu dem lokalen Achsenkreuz $\mathbf{e}(\alpha)$ in P bzw. P'. Die Differenz $\psi_a(P') - \psi_a(P) = d\psi_a$ ist das gewöhnliche Differential. Andererseits übertragen wir das Achsenkreuz $\mathfrak{e}(\alpha)$ von P nach P' durch Parallelverschiebung: $\mathfrak{e}'(\alpha)$; ψ'_a seien die Komponenten von ψ in P' in bezug auf das Achsenkreuz $\mathfrak{e}'(\alpha)$ daselbst. ψ_a wie ψ'_a hängen nur von der Wahl des Achsenkreuzes $\mathfrak{e}(\alpha)$ in P ab; sie haben nichts mit dem lokalen Achsenkreuz in P' zu tun. Mit der Drehung des Achsenkreuzes in P transformieren sich die ψ'_a ebenso wie die ψ_a, desgleichen die Differenzen $\delta\psi_a = \psi'_a - \psi_a$. Sie sind die Komponenten des kovarianten Differentials $\delta\psi$ von ψ. $\mathfrak{e}'(\alpha)$ geht aus dem lokalen Achsenkreuz $\mathfrak{e}(\alpha) = \mathfrak{e}(\alpha; P')$ in P' durch die in § 2 bestimmte infinitesimale Drehung $d\Omega$ hervor. Die entsprechende infinitesimale Transformation

$$dE = \tfrac{1}{2} d o(\beta\gamma) \,.\, A(\beta\gamma)$$

führt $\psi_a(P')$ in ψ'_a über, d. h. $\psi' - \psi(P')$ ist $= dE\,.\,\psi$. Addiert man $d\psi = \psi(P') - \psi(P)$, so erhält man

$$\delta\psi = d\psi + dE\,.\,\psi. \tag{13}$$

Alles hängt linear von der Verschiebung PP' ab. Es werde

$$\delta\psi = \psi_p (dx)^p = \psi(\alpha)\, v(\alpha), \quad dE = E_p (dx)^p = E(\alpha)\, v(\alpha)$$

geschrieben. Wir finden

$$\psi_p = \left(\frac{\partial}{\partial x_p} + E_p\right)\psi \quad \text{oder} \quad \psi(\alpha) = \left(e^p(\alpha)\,\frac{\partial}{\partial x_p} + E(\alpha)\right)\psi.$$

Darin ist

$$E(\alpha) = \tfrac{1}{2}\, o(\alpha;\, \beta\gamma)\, A(\beta\gamma).$$

Ist ψ' eine Größe von demselben Transformationsgesetz wie ψ, so sind

$$\psi^* S(\alpha)\, \psi'$$

die Komponenten eines Vektors mit Bezug auf das lokale Achsenkreuz. Darum ist

$$v'(\alpha) = \psi^* S(\alpha)\, \delta\psi = \psi^* S(\alpha)\, \psi(\beta)\,.\, v(\beta)$$

eine vom Achsenkreuz unabhängige lineare Abbildung $v \to v'$ des Vektorkörpers in P. Ihre Spur

$$\psi^* S(\alpha)\, \psi(\alpha)$$

ist folglich ein Skalar, und die Gleichung

$$i\,\varepsilon\,\mathfrak{m} = \psi^* S(\alpha)\, \psi(\alpha) \tag{14}$$

definiert eine skalare Dichte $\mathfrak{m}$, deren Integral

$$\int \mathfrak{m}\, dx \quad (dx = dx_0\, dx_1\, dx_2\, dx_3)$$

als Wirkungsgröße Verwendung finden kann.

Um zu einem expliziten Ausdruck von $\mathfrak{m}$ zu kommen, müssen wir

$$S(\alpha)\,E(\alpha) = \tfrac{1}{2}\,S(\alpha)\,A(\beta\,\gamma)\cdot o(\alpha;\,\beta\,\gamma) \tag{15}$$

ausrechnen. Aus (7) und (4) ergibt sich, daß

$$S(\beta)\,A(\beta\,\alpha) = \tfrac{1}{2}\,S(\alpha) \quad [\alpha \neq \beta, \text{ nicht über } \beta \text{ summieren!}]$$

ist und

$$S(\beta)\,A(\gamma\,\delta) = \tfrac{1}{2}\,S(\alpha),$$

wenn $\alpha\,\beta\,\gamma\,\delta$ eine gerade Permutation der Indizes 0 1 2 3 ist. Die Glieder der ersten und zweiten Art liefern darum als Beitrag zu (15) die folgenden Multipla von $S(\alpha)$:

$$\frac{1}{2}\,o(\varrho;\,\varrho\,\alpha) = \frac{1}{2\,\varepsilon}\,\frac{\partial\,\mathfrak{e}^p(\alpha)}{\partial x_p}$$

bzw.

$$o(\beta;\,\gamma\,\delta) + o(\gamma;\,\delta\,\beta) + o(\delta;\,\beta\,\gamma) = \frac{i}{2}\,\varphi(\alpha).$$

Nach (11) ist, wenn $\alpha\,\beta\,\gamma\,\delta$ eine gerade Permutation von 0 1 2 3 ist,

$$i\,\varphi(\alpha) = [\mathbf{e}(\beta),\ \mathbf{e}(\gamma)](\delta) + + \text{ (zykl. Permutationen von } \beta\,\gamma\,\delta)$$

$$= \sum \pm \frac{\partial\,e^p(\beta)}{\partial x_q}\,e^q(\gamma)\,e_p(\delta). \tag{16}$$

Die Summe erstreckt sich alternierend über die sechs Permutationen von $\beta\,\gamma\,\delta$ (außerdem natürlich über p und q). Mit diesen Bezeichnungen gilt

$$\mathfrak{m} = \frac{1}{i}\left(\psi^*\,\mathfrak{e}^p(\alpha)\,S(\alpha)\,\frac{\partial\psi}{\partial x_p} + \frac{1}{2}\,\frac{\partial\,\mathfrak{e}^p(\alpha)}{\partial x_p}\cdot\psi^*\,S(\alpha)\,\psi\right) + \frac{1}{4\,\varepsilon}\cdot\varphi(\alpha)\,\mathfrak{s}(\alpha). \tag{17}$$

Der zweite Teil ist

$$= \frac{1}{4\,i\,\varepsilon}\left|\,e_p(\alpha),\ e^q(\alpha),\ \frac{\partial\,e^p(\alpha)}{\partial x_q},\ \mathfrak{s}(\alpha)\,\right|$$

(summiert über p und q); jedes Glied ist eine Determinante von vier Zeilen, die man aus der hingeschriebenen Zeile erhält, wenn man der Reihe nach $\alpha = 0, 1, 2, 3$ setzt.

$$\mathfrak{s}(\alpha) \text{ ist } = \psi^*\,S(\alpha)\,\psi. \tag{18}$$

Nicht das Wirkungsintegral

$$\int \mathfrak{h}\,dx \tag{19}$$

selbst, sondern nur seine Variation ist von Bedeutung für die Naturgesetze. Darum ist es nicht nötig, daß $\mathfrak{h}$ reell ist, sondern es genügt, wenn die Differenz $\bar{\mathfrak{h}} - \mathfrak{h}$ eine Divergenz ist. In diesem Falle sagen wir, $\mathfrak{h}$ sei praktisch reell. Wir müssen prüfen, wie es in dieser Hinsicht mit $\mathfrak{m}$ bestellt ist. $e^p(\alpha)$ ist reell für $\alpha = 1, 2, 3$, rein imaginär für $\alpha = 0$. Darum ist $e^p(\alpha)\,S(\alpha)$ eine Hermitesche Matrix. Desgleichen ist $\varphi(\alpha)$ reell für $\alpha = 1, 2, 3$, rein imaginär für $\alpha = 0$; also ist auch $\varphi(\alpha)\,S(\alpha)$ hermitesch. Folglich ist

$$\overline{\mathfrak{m}} = -\frac{1}{i}\left(\frac{\partial \psi^*}{\partial x_p}\,\mathfrak{S}^p\,\psi + \frac{1}{2}\,\frac{\partial\, \mathfrak{e}^p(\alpha)}{\partial x_p}\cdot \psi^*\,S(\alpha)\,\psi\right) + \frac{1}{4\varepsilon}\cdot\varphi(\alpha)\,s(\alpha),$$

$$i\,(\mathfrak{m} - \overline{\mathfrak{m}}) = \psi^*\,\mathfrak{S}^p\,\frac{\partial \psi}{\partial x_p} + \frac{\partial \psi^*}{\partial x_p}\,\mathfrak{S}^p\,\psi + \frac{\partial\, \mathfrak{e}^p(\alpha)}{\partial x_p}\cdot\psi^*\,S(\alpha)\,\psi$$

$$= \frac{\partial}{\partial x_p}\,(\psi^*\,\mathfrak{S}^p\,\psi) = \frac{\partial\, \mathfrak{s}^p}{\partial x_p}\,.$$

$\mathfrak{m}$ ist also in der Tat praktisch reell.

Zur speziellen Relativitätstheorie kehren wir zurück, wenn wir

$$e^0(0) = -i, \quad e^1(1) = e^2(2) = e^3(3) = 1,$$

alle übrigen $e^p(\alpha) = 0$ setzen.

§ 4. Energie. (19) sei das Wirkungsintegral für die Materie im weiteren Sinne (Materie + elektrisches Feld), welche durch die ψ und die elektromagnetischen Potentiale f_p beschrieben ist. Die Naturgesetze drücken aus, daß die Variation

$$\delta\int \mathfrak{h}\,dx = 0$$

ist, wenn die ψ und f_p willkürlichen infinitesimalen Variationen unterworfen werden, die außerhalb eines endlichen Weltgebietes verschwinden. Die Variation der ψ gibt die materiellen Gleichungen im engeren Sinne, die Variation der f_p die elektromagnetischen Gleichungen. Auf Grund dieser Naturgesetze wird, wenn man auch die $e^p(\alpha)$, die bisher festgehalten wurden, einer analogen infinitesimalen Variation unterwirft, eine Gleichung bestehen

$$\delta\int \mathfrak{h}\,dx = \int \mathfrak{t}_p(\alpha)\,.\,\delta\,e^p(\alpha)\,.\,dx, \tag{20}$$

durch welche die Tensordichte $\mathfrak{t}_p(\alpha)$ der Energie zu definieren ist.

Zufolge der Invarianz der Wirkungsgröße muß (20) verschwinden, wenn die Variation $\delta\,e^p(\alpha)$ dadurch hervorgebracht wird,

1. daß bei festgehaltenem Koordinatensystem x_p das lokale Achsenkreuz $\mathbf{e}(\alpha)$ eine infinitesimale Drehung erleidet; oder
2. daß bei festgehaltenem Achsenkreuz die Koordinaten x_p einer infinitesimalen Transformation unterworfen werden.

Der erste Vorgang ist beschrieben durch die Gleichungen

$$\delta\,e^p(\alpha) = o(\alpha\beta)\,.\,e_p(\beta).$$

Hierin bilden die $o(\alpha\beta)$ eine schiefsymmetrische (infinitesimale) Matrix, die willkürlich vom Orte abhängt. Und das Verschwinden von (20) sagt aus, daß

$$t(\beta, \alpha) = t_p(\alpha)\, e^p(\beta)$$

symmetrisch ist in α und β. Die Symmetrie des Energietensors ist so mit der ersten Invarianzeigenschaft äquivalent. Das Symmetriegesetz ist aber nicht identisch erfüllt, sondern eine Folge der materiellen und elektromagnetischen Gesetze. Denn bei festgehaltenem ψ-Feld werden sich ja durch die Drehung des Achsenkreuzes die Komponenten von ψ ändern!

Etwas mühsamer ist die Berechnung der durch den zweiten Prozeß hervorgebrachten Variation $\delta e^p(\alpha)$. Aber die Überlegungen sind aus der Relativitätstheorie in ihrer früheren analytischen Fassung geläufig*. Der Punkt P mit den Koordinaten x_p habe im transformierten Koordinatensystem die Koordinaten

$$x'_p = x_p + \delta x_p, \quad \delta x_p = \xi^p(x).$$

Der Punkt, der im neuen Koordinatensystem dieselben Koordinaten x_p hat wie P im alten, werde mit P' bezeichnet; er hat im alten System die Koordinaten $x_p - \delta x_p$. Der Vektor $\boldsymbol{t}$ in P wird im neuen Koordinatensystem die Komponenten

$$\frac{\partial x'_p}{\partial x_q} \cdot t^q = t_p + \frac{\partial \xi^p}{\partial x_q} \cdot t^q$$

besitzen. Insbesondere ist die Änderung, welche die Komponenten $e^p(\alpha)$ des festen Vektors $\mathbf{e}(\alpha)$ im festgehaltenen Punkt P durch die Koordinatentransformation erleiden,

$$\delta' e^p(\alpha) = \frac{\partial \xi^p}{\partial x_q} \cdot e^q(\alpha).$$

Andererseits ist der Unterschied zwischen dem Vektor $\mathbf{e}(\alpha)$ in P' und P gegeben durch

$$d e^p(\alpha) = -\frac{\partial e^p(\alpha)}{\partial x_q} \cdot \xi^q.$$

Darum ist die Variation, welche durch die Koordinatentransformation bei festgehaltenen Koordinatenwerten x_p erzeugt wird:

$$\delta e^p(\alpha) = \frac{\partial \xi^p}{\partial x_q} \cdot e^q(\alpha) - \frac{\partial e^p(\alpha)}{\partial x_q} \cdot \xi^q.$$

---

* Vgl. etwa H. Weyl, Raum, Zeit, Materie, 5. Aufl., S. 233 ff. (zitiert als RZM). Berlin 1923.

Hier sind ξ^p willkürliche, außerhalb eines endlichen Weltgebietes verschwindende Funktionen. Setzen wir in (20) ein, so erhalten wir durch eine partielle Integration

$$0 = \int \left\{ \frac{\partial \mathfrak{t}_p^q}{\partial x_q} + \mathfrak{t}_q(\alpha) \frac{\partial e^q(\alpha)}{\partial x_p} \right\} \xi^p \, dx.$$

Der Quasi-Erhaltungssatz von Energie und Impuls ergibt sich demnach hier in der Gestalt

$$\frac{\partial \mathfrak{t}_p^q}{\partial x_q} + \frac{\partial e^q(\alpha)}{\partial x_p} \mathfrak{t}_q(\alpha) = 0. \tag{21}$$

Wegen des zweiten Gliedes ist er nur in der speziellen Relativitätstheorie ein wirklicher Erhaltungssatz. In der allgemeinen wird er erst dazu wenn die Energie des Gravitationsfeldes hinzugefügt wird.

In der speziellen Relativitätstheorie aber liefert Integration nach $d\xi = dx_1 \, dx_2 \, dx_3$ über den räumlichen Querschnitt

$$x_0 = t = \text{const} \tag{22}$$

die zeitlich konstanten Komponenten von Impuls (J_1, J_2, J_3) und Energie ($-J_0$):

$$J_p = \int \mathfrak{t}_p^0 \, d\xi.$$

Mit Hilfe der Symmetrie findet man ferner die Divergenzgleichungen

$$\frac{\partial}{\partial x_q}(x_2 \mathfrak{t}_3^q - x_3 \mathfrak{t}_2^q) = 0, \ldots,$$

$$\frac{\partial}{\partial x_q}(x_0 \mathfrak{t}_1^q + x_1 \mathfrak{t}_0^q) = 0, \ldots$$

Die drei Gleichungen der ersten Art zeigen, daß das **Impulsmoment** (M_1, M_2, M_3) zeitlich konstant ist:

$$M_1 = \int (x_2 \mathfrak{t}_3^0 - x_3 \mathfrak{t}_2^0) \, d\xi, \ldots,$$

die Gleichungen der zweiten Art enthalten den Satz von der **Trägheit der Energie**.

Wir rechnen die Energiedichte aus für die oben aufgestellte Wirkungsgröße $\mathfrak{m}$ der Materie; wir behandeln die beiden Teile, in die $\mathfrak{m}$ nach (17) zerlegt erscheint, gesondert. Für den ersten Teil kommt nach einer partiellen Integration

$$\int \delta \mathfrak{m} \, . \, dx = \int \mathfrak{u}_p(\alpha) \, \delta \mathfrak{e}^p(\alpha) \, . \, dx$$

mit

$$i\, \mathfrak{u}_p(\alpha) = \psi^* S(\alpha) \frac{\partial \psi}{\partial x_p} - \frac{1}{2} \frac{\partial (\psi^* S(\alpha) \psi)}{\partial x_p},$$

$$= \frac{1}{2}\left(\psi^* S(\alpha) \frac{\partial \psi}{\partial x_p} - \frac{\partial \psi^*}{\partial x_p} S(\alpha) \psi \right).$$

Der hieraus entspringende Teil der Energie ist darum

$$\mathfrak{t}_p(\alpha) = \mathfrak{u}_p(\alpha) - e_p(\alpha) \cdot \mathfrak{u}, \quad \mathfrak{t}_p^q = \mathfrak{u}_p^q - \delta_p^q \mathfrak{u},$$

wo $\mathfrak{u}$ die Verkürzung $e^p(\alpha)\,\mathfrak{u}_p(\alpha)$ bedeutet. Diese Formeln sind allgemein auch für nichtkonstante $e^p(\alpha)$ richtig. Im zweiten Teil beschränken wir uns aber der Einfachheit halber auf die spezielle Relativität. Für ihn ist dann

$$\int \delta\,\mathfrak{m} \, . \, dx = \frac{1}{4i} \int \left| e_p(\alpha),\; e^q(\alpha),\; \frac{\partial(\delta\, e^p(\alpha))}{\partial x_q},\; \mathfrak{s}(\alpha) \right| dx,$$

$$= -\frac{1}{4i} \int \left| \delta\, e^p(\alpha),\; e_p(\alpha),\; e^q(\alpha),\; \frac{\partial\, \mathfrak{s}(\alpha)}{\partial x_q} \right| dx,$$

$$\mathfrak{t}_p(0) = -\frac{1}{4i} \left| e_p(\alpha),\; e^q(\alpha),\; \frac{\partial\, \mathfrak{s}(\alpha)}{\partial x_q} \right|_{\alpha = 1, 2, 3}.$$

$\mathfrak{t}_p^0$ entsteht daraus durch Multiplikation mit $-i$; somit $\mathfrak{t}_0^0 = 0$ und

$$\mathfrak{t}_1^0 = \frac{1}{4}\left(\frac{\partial\,\mathfrak{s}(3)}{\partial x_2} - \frac{\partial\,\mathfrak{s}(2)}{\partial x_3}\right). \tag{23}$$

Wir vereinigen beide Bestandteile, um totale Energie, Impuls und Impulsmoment zu bestimmen. Aus

$$\mathfrak{t}_0^0 = -\frac{1}{2i} \sum_{p=1}^{3} \left(\psi^* S^p \frac{\partial \psi}{\partial x_p} - \frac{\partial \psi^*}{\partial x_p} S^p \psi\right)$$

ergibt sich nach einer auf den Subtrahenden ausgeübten partiellen Integration

$$-J_0 = -\int \mathfrak{t}_0^0 \, d\xi = \frac{1}{i} \int \psi^* \cdot \sum_{p=1}^{3} S^p \frac{\partial \psi}{\partial x_p} \cdot d\xi.$$

Dies führt dazu zurück, den Operator

$$\frac{1}{i} \sum_{p=1}^{3} S^p \frac{\partial}{\partial x_p}$$

als Repräsentanten der Energie einer freien Partikel anzusetzen. Ferner wird

$$J_1 = \int \mathfrak{t}_1^0 \, d\xi = \frac{1}{2i} \int \left(\psi^* \frac{\partial \psi}{\partial x_1} - \frac{\partial \psi^*}{\partial x_1} \psi\right) d\xi$$

$$= \frac{1}{i} \int \psi^* \frac{\partial \psi}{\partial x_1} \, d\xi.$$

Das Glied (23) liefert zum Integral keinen Beitrag. Der Impuls wird, wie es nach Schrödinger sein muß, durch den Operator

$$\frac{1}{i}\left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}\right)$$

dargestellt. Aus dem vollständigen Ausdruck von

$$x_2 \mathfrak{t}_3^0 - x_3 \mathfrak{t}_2^0$$

erhält man schließlich durch geeignete partielle Integrationen

$$M_1 = \int \left\{ \frac{1}{i} \psi^* \left(x_2 \frac{\partial \psi}{\partial x_3} - x_3 \frac{\partial \psi}{\partial x_2} \right) + \frac{1}{2} s(1) \right\} d\xi.$$

Im Einklang mit bekannten Formeln ist also M_1 repräsentiert durch den Operator

$$\frac{1}{i} \left(x_2 \frac{\partial}{\partial x_3} - x_3 \frac{\partial}{\partial x_2} \right) + \frac{1}{2} S(1).$$

Nachdem man den Spin von Anbeginn in die Theorie hineingesteckt hat, muß er natürlich hier wieder zum Vorschein kommen; es ist aber doch recht überraschend und instruktiv, wie das zustande kommt. Die grundlegenden Ansätze der Quantentheorie haben hiernach einen weniger prinzipiellen Charakter, als man wohl ursprünglich angenommen hatte. Sie sind an die spezielle Wirkungsgröße $\mathfrak{m}$ gebunden. Andererseits bestätigt dieser Zusammenhang die Unersetzbarkeit von $\mathfrak{m}$ in seiner Rolle als Wirkung der Materie. Nur die allgemeine Relativitätstheorie, die durch ihre freie Veränderlichkeit der $e^p(\alpha)$ zu einer willkürfreien Definition der Energie führt, erlaubt uns, in der geschilderten Weise den Zirkel der Quantentheorie zu schließen.

§ 5. Gravitation. Wir nehmen die Transkription von Einsteins klassischer Gravitationstheorie wieder auf und bestimmen zunächst den Riemannschen Krümmungstensor*. Von dem Punkte P führen die Linienelemente d und δ nach P_d und P_δ. Das Linienelement δ wird irgendwie nach P_d, d nach P_δ so überführt, daß sie sich in einer gemeinsamen, P gegenüberliegenden Ecke P^* eines infinitesimalen „Parallelogramms" treffen. Das Achsenkreuz $\mathfrak{e}(\alpha)$ in P wird einmal auf dem Wege $P P_d P^*$, ein andermal auf dem Wege $P P_\delta P^*$ nach P^* parallel übertragen. Die beiden normalen Achsenkreuze, die man so in P^* erhält, gehen durch eine infinitesimale Drehung

$$P_{pq}(dx)^p(\delta x)^q = \tfrac{1}{2} P_{pq}(\varDelta x)^{pq}$$

auseinander hervor, wo

$$(\varDelta x)^{pq} = (dx)^p(\delta x)^q - (\delta x)^p(dx)^q$$

die Komponenten des von dx und δx aufgespannten Flächenelements sind und P_{pq} schiefsymmetrisch ist in bezug auf p und q. P_{pq} ist eine schiefsymmetrische Matrix $\| r_{pq}(\alpha\beta) \|$; das ist der Riemannsche Krümmungstensor.

* Vgl. RZM, S. 119f.

Die Drehung, welche das auf dem ersten Wege nach P^* parallel verschobene Achsenkreuz $\mathfrak{e}^*(\alpha)$ aus dem lokalen Achsenkreuz $\mathfrak{e}(\alpha)$ in P^* erzeugt, ist in einer leicht verständlichen Bezeichnung

$$(1 + d\,\Omega)\bigl(1 + \delta\,\Omega(P_d)\bigr).$$

Die Differenz dieses Ausdrucks und desjenigen, der daraus durch Vertauschung von d und δ hervorgeht, ist

$$= \{d(\delta\,\Omega) - \delta(d\,\Omega)\} + (d\,\Omega\,.\,\delta\,\Omega - \delta\,\Omega\,.\,d\,\Omega).$$

$$d\,\Omega \text{ ist } = \Omega_p\,(d\,x)^p,$$

$$\delta\,(d\,\Omega) = \frac{\partial\,\Omega_p}{\partial\,x_q}\,\delta\,x_q\,d\,x_p + \Omega_p\,\delta\,d\,x_p.$$

Weil das Parallelogramm sich schließt, ist $\delta\,d\,x_p = d\,\delta\,x_p$; darum schließlich

$$P_{p\,q} = \left(\frac{\partial\,\Omega_q}{\partial\,x_p} - \frac{\partial\,\Omega_p}{\partial\,x_q}\right) + (\Omega_p\,\Omega_q - \Omega_q\,\Omega_p).$$

Zur skalaren Krümmung

$$r = e^p(\alpha)\,e^q(\beta)\,r_{p\,q}(\alpha\,\beta)$$

liefert der erste, differentiierte Bestandteil den Beitrag

$$\bigl(e^q(\alpha)\,e^p(\beta) - e^q(\beta)\,e^p(\alpha)\bigr)\,\frac{\partial\,o_p(\alpha\,\beta)}{\partial\,x_q}.$$

In $\mathfrak{r} = \dfrac{r}{\varepsilon}$ liefert er, unter Vernachlässigung einer vollständigen Divergenz, die beiden Glieder

$$-\,2\,o(\beta,\,\alpha\,\beta)\,\frac{\partial\,\mathfrak{e}^q(\alpha)}{\partial\,x_q}$$

und

$$\frac{1}{\varepsilon}\,o_p(\alpha\,\beta)\left\{\frac{\partial\,e^p(\alpha)}{\partial\,x_q}\,e^q(\beta) - \frac{\partial\,e^p(\beta)}{\partial\,x_q}\,e^q(\alpha)\right\}.$$

Das erste ist nach (12)

$$= -\,2\,o(\beta;\,\varrho\,\beta)\,o(\alpha;\,\alpha\,\varrho),$$

das zweite nach (10)

$$= 2\,o(\alpha;\,\beta\,\gamma)\,.\,o(\gamma;\,\alpha\,\beta).$$

Das Resultat ist der folgende Ausdruck für die Wirkungsdichte $\mathfrak{g}$ der Gravitation

$$\varepsilon\,\mathfrak{g} = o(\alpha;\,\beta\,\gamma)\,.\,o(\gamma;\,\alpha\,\beta) + o(\alpha;\,\alpha\,\gamma)\,.\,o(\beta;\,\beta\,\gamma). \qquad (24)$$

Das Integral $\int \mathfrak{g}\,d\,x$ ist nicht wirklich, aber praktisch invariant, $\mathfrak{g}$ unterscheidet sich von der skalaren Dichte $\mathfrak{r}$ um eine Divergenz.

Variation der $e^p(\alpha)$ in dem totalen Wirkungsintegral

$$\int (\mathfrak{g} + \varkappa\,\mathfrak{h})\,d\,x$$

liefert die Gravitationsgleichungen ($\varkappa$ ist eine numerische Konstante).

Die Gravitationsenergie $\mathfrak{v}_p^q$ erhält man aus $\mathfrak{g}$, wenn man im Koordinatenraum eine infinitesimale Verschiebung vornimmt*:

$$x'_p = x_p + \xi^p, \quad \xi^p = \text{const.}$$

Die dadurch hervorgebrachte Variation

$$\delta\, \mathbf{e}(\alpha) \text{ ist } = -\frac{\partial\, \mathbf{e}(\alpha)}{\partial x_p} \cdot \xi^p.$$

$\mathfrak{g}$ ist eine Funktion von $e^p(\alpha)$ und den Ableitungen $e_q^p(\alpha) = \dfrac{\partial e^p(\alpha)}{\partial x_q}$; das totale Differential sei mit

$$\delta\, \mathfrak{g} = \mathfrak{g}_p(\alpha)\, \delta\, e^p(\alpha) + \mathfrak{g}_p^q(\alpha)\, \delta\, e_q^p(\alpha)$$

bezeichnet. Für die durch die infinitesimale Translation im Koordinatenraum hervorgerufene Änderung muß

$$\int \delta\, \mathfrak{g} \cdot dx + \int \frac{\partial\, \mathfrak{g}}{\partial x_p}\, \xi^p \cdot dx = 0 \qquad (25)$$

gelten; das Integral erstreckt sich über ein beliebiges Stück der Welt.

$$\int \delta\, \mathfrak{g} \cdot dx = \int \left(\mathfrak{g}_p(\alpha) - \frac{\partial\, \mathfrak{g}_p^q(\alpha)}{\partial x_q} \right) \delta\, e^p(\alpha) \cdot dx + \int \frac{\partial \left(\mathfrak{g}_p^q(\alpha)\, \delta\, e^p(\alpha) \right)}{\partial x_q}\, dx.$$

Nach den Gravitationsgleichungen ist die Klammer im ersten Integral $= -\varkappa\, \mathfrak{t}_p(\alpha)$, das Integral selbst

$$= -\varkappa \int \mathfrak{t}_q(\alpha)\, \frac{\partial e^q(\alpha)}{\partial x_p} \cdot \xi^p\, dx.$$

Man führe

$$\mathfrak{v}_p^q = \delta_p^q\, \mathfrak{g} - \frac{\partial e^r(\alpha)}{\partial x_p} \cdot \mathfrak{g}_r^q(\alpha)$$

ein. Gleichung (25) besagt, daß das über ein beliebiges Weltstück erstreckte Integral von

$$\left(\mathfrak{v}_p^q - \varkappa\, \mathfrak{t}_q(\alpha)\, \frac{\partial e^q(\alpha)}{\partial x_p} \right) \xi^p$$

Null ist. Der Integrand muß darum überall verschwinden. Da die ξ^p beliebige Konstanten sind, sind die Faktoren von ξ^p einzeln Null. So verwandelt sich (21) in die reine Divergenzgleichung

$$\frac{\partial \left(\mathfrak{v}_p^q + \varkappa\, \mathfrak{t}_p^q \right)}{\partial x_q} = 0,$$

und $\mathfrak{v}_p^q/\varkappa$ erweist sich als Gravitationsenergie.

Um daneben einen wirklichen differentiellen Erhaltungssatz des Impulsmoments in der allgemeinen Relativitätstheorie formulieren zu

---

* Vgl. RZM, S. 272f.

können, muß man die Koordinaten so spezialisieren, daß die kongrediente Drehung aller Achsenkreuze als eine orthogonale Transformation der Koordinaten erscheint. Dies ist sicher möglich, doch gehe ich hier darauf nicht näher ein.

§ 6. Elektrisches Feld. Wir kommen jetzt zu dem kritischen Teil der Theorie. Meiner Meinung nach liegt der Ursprung und die Notwendigkeit des elektromagnetischen Feldes in folgendem begründet. Die Komponenten ψ_1, ψ_2 sind in Wahrheit nicht eindeutig durch das Achsenkreuz bestimmt, sondern nur insoweit, daß sie noch mit einem beliebigen „Eichfaktor" $e^{i\lambda}$ vom absoluten Betrag 1 multipliziert werden können. Nur bis auf einen solchen Faktor ist die Transformation bestimmt, welche die ψ unter dem Einfluß einer Drehung des Achsenkreuzes erleiden. In der speziellen Relativitätstheorie muß man diesen Eichfaktor als eine Konstante ansehen, weil wir hier ein einziges, nicht an einen Punkt gebundes Achsenkreuz haben. Anders in der allgemeinen Relativitätstheorie: jeder Punkt hat sein eigenes Achsenkreuz und darum auch seinen eigenen willkürlichen Eichfaktor; dadurch, daß man die starre Bindung der Achsenkreuze in verschiedenen Punkten aufhebt, wird der Eichfaktor notwendig zu einer willkürlichen Ortsfunktion. Dann ist aber auch die infinitesimale lineare Transformation dE der ψ, welche der infinitesimalen Drehung $d\Omega$ entspricht, nicht vollständig festgelegt, sondern dE kann um ein beliebiges rein imaginäres Multiplum $i \,.\, df$ der Einheitsmatrix vermehrt werden. Zur eindeutigen Festlegung des kovarianten Differentials $\delta\psi$ von ψ hat man also außer der Metrik in der Umgebung des Punktes P ein solches df für jedes von P ausgehende Linienelement $\overrightarrow{PP'} = (dx)$ nötig. Damit $\delta\psi$ nach wie vor linear von dx abhängt, muß

$$df = f_p\,(dx)^p$$

eine Linearform in den Komponenten des Linienelements sein. Ersetzt man ψ durch $e^{i\lambda} . \psi$, so muß man zugleich, wie aus der Formel für das kovariante Differential hervorgeht, df ersetzen durch $df - d\lambda$.

Dies hat zur Folge, daß zur Wirkungsdichte $\mathfrak{m}$ das Glied

$$\frac{1}{\varepsilon} f(\alpha)\, s(\alpha) = \frac{1}{\varepsilon} f(\alpha) \,.\, \psi^* S(\alpha)\, \psi = f_p \,.\, \psi^* \mathfrak{S}^p\, \psi \tag{26}$$

zu addieren ist. $\mathfrak{m}$ bedeutet fortan die so ergänzte Wirkungsgröße. Es herrscht notwendig Eichinvarianz in dem Sinne, daß die Wirkungsgröße ungeändert bleibt, wenn

$$\psi \text{ durch } e^{i\lambda} . \psi, \quad f_p \text{ durch } f_p - \frac{\partial \lambda}{\partial x_p}$$

ersetzt wird, unter λ eine willkürliche Ortsfunktion verstanden. Genau in der durch (26) beschriebenen Weise wirkt nach der Erfahrung das elektromagnetische Potential auf die Materie. Wir sind daher berechtigt, die hier eingeführten Größen f_p mit den Komponenten jenes Potentials zu identifizieren. Der Beweis ist vollkommen, wenn wir zeigen, daß auch umgekehrt das f_p-Feld nach denselben Gesetzen von der Materie beeinflußt wird, welche nach der Erfahrung für das elektromagnetische Potentialfeld gelten.

$$f_{pq} = \frac{\partial f_q}{\partial x_p} - \frac{\partial f_p}{\partial x_q}$$

ist ein eichinvarianter schiefsymmetrischer Tensor und

$$\mathfrak{l} = \frac{1}{4} f_{pq} \mathfrak{f}^{pq} \tag{27}$$

die für die Maxwellsche Theorie charakteristische skalare Dichte. Der Ansatz

$$\mathfrak{h} = \mathfrak{m} + a\,\mathfrak{l} \tag{28}$$

(a eine numerische Konstante) liefert durch Variation der f_p die Maxwellschen Gleichungen mit

$$-\mathfrak{s}^p = -\psi^* \mathfrak{S}^p \psi \tag{29}$$

als der Dichte des elektrischen Viererstroms.

Die Eichinvarianz steht in engem Zusammenhang mit dem Erhaltungssatz für die Elektrizität. Weil $\mathfrak{h}$ eichinvariant ist, muß $\delta \int \mathfrak{h}\, dx$ identisch verschwinden, wenn bei festgehaltenen $e^p(\alpha)$ die ψ und f_p gemäß

$$\delta \psi = i\lambda \cdot \psi, \quad \delta f_p = -\frac{\partial \lambda}{\partial x_p}$$

variiert werden; λ ist eine willkürliche Ortsfunktion. Dies liefert eine identisch erfüllte Relation zwischen den materiellen und den elektromagnetischen Gleichungen. Wissen wir, daß die materiellen Gleichungen (im engeren Sinne) gelten, so folgt also daraus

$$\delta \int \mathfrak{h}\, dx = 0,$$

wenn nur die f_p gemäß der Gleichung $\delta f_p = -\partial \lambda / \partial x_p$ variiert werden. Andererseits folgt aus den elektromagnetischen Gleichungen dasselbe für die infinitesimale Variation $\delta \psi = i \lambda \cdot \psi$ der ψ allein. Wenn $\mathfrak{h} = \mathfrak{m} + a\,\mathfrak{l}$, erhält man beide Male

$$\int \delta \mathfrak{h} \cdot dx = \pm \int \psi^* \mathfrak{S}^p \psi \cdot \frac{\partial \lambda}{\partial x_p}\, dx = \mp \int \lambda \frac{\partial \mathfrak{s}^p}{\partial x_p}\, dx.$$

Eine analoge Sachlage fanden wir vor beim Erhaltungssatz von Energieimpuls und Impulsmoment. Sie verknüpften die materiellen Gleichungen im weiteren Sinne mit den Gravitationsgleichungen und korrespondierten der Invarianz gegenüber Koordinatentransformationen bzw. gegenüber willkürlichen unabhängigen Drehungen der lokalen Achsenkreuze in den verschiedenen Weltpunkten.

Aus
$$\frac{\partial \mathfrak{s}^p}{\partial x_p} = 0 \tag{30}$$
ergibt sich, daß der Fluß der Vektordichte $\mathfrak{s}^p$ durch einen dreidimensionalen Querschnitt der Welt, insbesondere durch einen Querschnitt (22)
$$l = \int \mathfrak{s}^0 \, d\xi \tag{31}$$
von der Lage des Querschnitts bzw. von t unabhängig ist. Nicht nur dieses Integral, sondern auch das einzelne Integralelement hat eine invariante Bedeutung; immerhin hängt das Vorzeichen davon ab, welcher Richtungssinn als eine positive Überquerung des dreidimensionalen Schnitts gerechnet wird. Um $\mathfrak{s}^0 d\xi$ als räumliche Wahrscheinlichkeitsdichte ansprechen zu können, muß die Hermitesche Form
$$e^0(\alpha) \, . \, \psi^* S(\alpha) \, \psi \tag{32}$$
von ψ_1, ψ_2 definit sein. Man findet leicht, daß dies dann der Fall ist, wenn (22) wirklich ein räumlicher Querschnitt in P ist, wenn die in ihm liegenden, von P ausgehenden Linienelemente raumartig sind. Damit (32) das positive Vorzeichen bekommt, müssen die Querschnitte $x_0 =$ const, nach wachsendem x_0 geordnet, in der durch den Vektor $\mathbf{e}(0)/i$ angezeigten Richtung der Zukunft aufeinanderfolgen. Unter diesen sich hier naturgemäß ergebenden Einschränkungen des Koordinatensystems ist auch das Vorzeichen des Flusses festgelegt, und die Invariante (31) werde in der üblichen Weise durch die Bedingung
$$l \equiv \int \mathfrak{s}^0 \, d\xi = 1 \tag{33}$$
normiert. Die $\mathfrak{m}$ und $\mathfrak{l}$ miteinander kombinierende Konstante a ist dann eine reine Zahl $= ch/e^2$ (reziproke Feinstrukturkonstante).

Wir behandeln ψ_1, ψ_2; f_p; $e^p(\alpha)$ als die unabhängig voneinander zu variierenden Größen. Die aus $\mathfrak{m}$ entspringende Energiedichte $\mathfrak{t}_p^q$ muß wegen des Ergänzungsgliedes (26) um
$$f_p \mathfrak{s}^q - \delta_p^q (f_r \mathfrak{s}^r)$$
vermehrt werden. Das führt dazu, in der speziellen Relativitätstheorie der Energie den Operator
$$H = \sum_{p=1}^{3} S^p \left(\frac{1}{i} \frac{\partial}{\partial x_p} + f_p \right)$$
zuzuordnen, da
$$\int \psi^* . H \psi . d\xi$$

ihr Wert ist. Die materiellen Gleichungen lauten dann freilich

$$\left(\frac{1}{i}\frac{\partial}{\partial x_0} + f_0\right)\psi + H\psi = 0 \quad \text{und nicht} \quad \frac{1}{i}\frac{\partial \psi}{\partial x_0} + H\psi = 0,$$

wie es bisher in der Quantenmechanik angenommen wurde. Natürlich tritt zu der materiellen Energie die elektromagnetische hinzu, für welche die klassischen Ausdrücke Maxwells ihre Gültigkeit behalten.

Was die physikalischen Dimensionen betrifft, so ist es bei allgemeiner Relativität natürlich, die Koordinaten x_p als reine Zahlen anzusetzen. Die auftretenden Größen sind nicht nur invariant gegenüber Maßstabsänderung, sondern gegenüber beliebigen Transformationen der x_p. Werden alle $\mathfrak{e}(\alpha)$ durch Multiplikation mit einer Konstante b in $b\,.\,\mathfrak{e}(\alpha)$ verwandelt, so muß, wenn die Normierung (33) aufrechterhalten wird, gleichzeitig ψ durch $b^{3/2}\,.\,\psi$ ersetzt werden. $\mathfrak{m}$ und $\mathfrak{l}$ werden dadurch nicht verändert, sind also reine Zahlen. Dagegen nimmt $\mathfrak{g}$ den Faktor $1/b^2$ an, so daß $\varkappa$ das Quadrat einer Länge d ist. $\varkappa$ ist nicht identisch mit der Einsteinschen Gravitationskonstante, sondern entsteht aus ihr durch Multiplikation mit $2\,h/c$. d liegt weit unter der atomaren Größenordnung, es ist $\sim 10^{-32}$ cm. So wird auch hier die Gravitation nur für die astronomischen Probleme von Bedeutung sein.

Sehen wir vom Gravitationsglied ab, so enthalten die Feldgleichungen keine dimensionierte Atomkonstante. Für eine Wirkungsgröße wie das Glied in der Diracschen Theorie, das die Masse als Faktor trägt*, ist in der Zweikomponententheorie kein Platz. Aber man weiß, wie auf Grund der Erhaltungssätze die Masse eingeführt werden kann. Man nehme an, daß in der „leeren Umwelt" des Teilchens, außerhalb eines gewissen Weltkanals, dessen Querschnitte $x_0 = \text{const}$ von endlicher Ausdehnung sind, die $\mathfrak{t}_p^q$ verschwinden und die $e^p(\alpha)$ die konstanten Werte der speziellen Relativität annehmen. Dann sind

$$J_p = \int \left(\mathfrak{t}_p^0 + \frac{1}{\varkappa}\mathfrak{v}_p^0\right) d\xi$$

die Komponenten eines von der Willkür des Koordinatensystems und der lokalen Achsenkreuze nicht beeinflußten, zeitlich konstanten Vierervektors in der Umwelt. Das normale Koordinatensystem daselbst kann genauer festgelegt werden durch die Bedingung, daß der Impuls (J_1, J_2, J_3) verschwindet; dann ist $-J_0$ die invariante und zugleich konstante Masse des Teilchens. Es werde nun verlangt, daß diese Masse einen ein für allemal vorgegebenen Wert m hat.

* Proc. Roy. Soc. (A) **117**, 610.

Neben der hier besprochenen Theorie des elektromagnetischen Feldes, die ich für die richtige halte, weil sie so natürlich aus der Willkürlichkeit des Eichfaktors in ψ entspringt und darum die erfahrungsgemäß bestehende Eichinvarianz in Zusammenhang mit dem Erhaltungssatz für die Elektrizität verstehen läßt, bietet sich noch eine andere dar, welche die Elektrizität mit der Gravitation verknüpft. Das Glied (26) hat die gleiche Form wie der zweite Teil von $\mathfrak{m}$, Formel (17); $\varphi(\alpha)$ spielt in diesem die gleiche Rolle wie $f(\alpha)$ in jenem. Man mag daher erwarten, daß Materie und Gravitation, ψ und $e^p(\alpha)$, für sich schon ausreichen, die elektromagnetischen Erscheinungen zu erklären, indem man die Größen $\varphi(\alpha)$ als die elektromagnetischen Potentiale anspricht. Jene Größen hängen so von den $e^p(\alpha)$ und ihren ersten Ableitungen ab, daß Invarianz stattfindet gegenüber beliebigen Koordinatentransformationen. Was aber die Drehungen der Achsenkreuze anlangt, so transformieren sich die $\varphi(\alpha)$ nur dann wie die Komponenten eines festen Vektors im Achsenkreuz, wenn die Achsenkreuze in allen Punkten derselben Drehung unterworfen werden. Ignoriert man das Materiefeld und achtet nur auf den Zusammenhang von Elektrizität und Gravitation, so kommt man also auf eine Theorie der Elektrizität genau von der gleichen Art, wie sie Einstein neuerdings versucht hat. Immerhin wäre hier der Fernparallelismus nur vorgetäuscht.

Ich habe mich überzeugt, daß man durch diesen vielleicht zunächst verlockenden Ansatz nicht zu den Maxwellschen Gleichungen gelangt. Ferner bliebe die Eichinvarianz ganz rätselhaft; das elektromagnetische Potential selbst und nicht bloß die Feldstärke hätte physikalische Bedeutung. Darum glaube ich, daß diese Idee auf einen Holzweg führt, daß wir vielmehr dem Fingerzeig zu trauen haben, den uns die Eichinvarianz gab: die Elektrizität ein Begleitphänomen des materiellen Wellenfeldes und nicht der Gravitation.

Palmer Physical Laboratory, Princeton University, 19. April 1929.

Chapter 13
Erwin Schrödinger (1932): Diracsches Elektron im Schwerefeld I

Erwin Schrödinger (1932). Diracsches Elektron im Schwerefeld I. *Sitzungsberichte der preussischen Akademie der Wissenschaften, physikalisch-mathematische Klasse*, 1932, 105–128.

Diracsches Elektron im Schwerefeld I.

Von E. Schrödinger.

(Vorgelegt am 25. Februar 1932 [s. oben S. 46].)

§ 1. Einleitung.

Die Vereinigung der Diracschen Theorie des Elektrons mit der allgemeinen Relativitätstheorie ist schon wiederholt in Angriff genommen worden, so von Wigner[1], Tetrode[2], Fock[3], Weyl[4], Zaycoff[5], Podolsky[6]. Die meisten Autoren führen in jedem Weltpunkt ein orthogonales Achsenkreuz und in bezug auf dieses numerisch spezialisierte Diracsche Matrizen ein. Bei diesem Verfahren ist es ein bißchen schwer, zu erkennen, ob die Einsteinsche Idee des Fernparallelismus, auf die teilweise direkt Bezug genommen wird, wirklich hereinspielt oder ob man davon unabhängig ist. Ferner wird es dabei nötig, die Riemanngeometrischen Begriffe in die weniger vertraute und entschieden umständlichere Form der »Beinkomponenten« umzugießen. Es schien mir wünschenswert, all dies zu vermeiden, indem man, wie Tetrode[7], nur auf die verallgemeinerten Vertauschungsrelationen [s. u. Gleichung (2)] sich stützt. Es zeigt sich, daß man so außerordentlich einfach und geradlinig zu den wichtigen Operatoren Γ_k geführt wird, deren Spur das Viererpotential ist und die Fock als »Komponenten der Parallelverschiebung eines Spinors« einführt; und ebenso geradlinig zu dem wichtigen Gleichungssystem [s. u. (8)], das Fock auf dem Umweg über die Beinkomponenten gewinnt. Durch eine Beschränkung der zulässigen Bezugsysteme (s. u. § 4), welche der in der speziellen Relativitätstheorie üblichen völlig analog ist, führt man sodann die für die Interpretation wünschenswerten Hermitizitäten herbei, sowie eine Zuordnung zwischen Tensoroperatoren und lokalen c-Tensoren, welche ebenfalls völlig analog ist der in der speziellen Theorie von v. Neumann[8] aufgestellten [s. Gleichung (57) unten]. Ein grundsätzlicher Vorzug scheint es mir, daß sich der ganze Apparat fast vollständig durch reinen Operatorkalkül aufbauen läßt, ohne auf die ψ-Funktion Bezug zu nehmen. Hoffentlich

---

[1] E. Wigner, ZS. f. Phys. **53**, 592, 1928.
[2] H. Tetrode, ZS. f. Phys. **50**, 336, 1928.
[3] V. Fock, ZS. f. Phys. **57**, 261, 1929.
[4] H. Weyl, Proc. Nat. Acad. Amer. **15**, 323, 1929; ZS. f. Phys. **56**, 330, 1929.
[5] R. Zaycoff, Ann. d. Phys. **7**, 650, 1930.
[6] B. Podolsky, Phys. Rev. **37**, 1398, 1931.
[7] So auch B. Hoffmann in einem Brief an Phys. Rev. **37**, 88, 1931.
[8] J. v. Neumann, ZS. f. Phys. **48**, 868, 1928.

erschreckt die exakte Begründung dieses Apparats nicht allzusehr durch ihren Umfang, an dem die breite Schreibweise des Autors mitschuldig ist. Nach dieser einmal geleisteten Vorarbeit dürfte sich die Handhabung und das Denken damit als einfach erweisen. — Meine ausgiebige Verschuldung an die Arbeiten der Vorgänger möchte ich hier ein für allemal erklären, bitte aber aus methodischen Gründen, alles neu herleiten zu dürfen, so als wäre es noch von niemandem gefunden.

§ 2. Aufbau der Metrik aus Matrizenfeldern.

Wir nennen die Weltvariablen

$$x_0 = ict\,, \quad x_1 = x\,, \quad x_2 = y\,, \quad x_3 = z\,.$$

Die erste ist stets reinimaginär, die anderen drei reell. Diracs Grundgedanke war, den euklidischen Wellenoperator

$$\frac{\partial^2}{\partial x_0^2} + \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2}$$

als das Quadrat eines linearen Operators aufzufassen

$$\left(\mathring{\gamma}_0 \frac{\partial}{\partial x_0} + \mathring{\gamma}_1 \frac{\partial}{\partial x_1} + \mathring{\gamma}_2 \frac{\partial}{\partial x_2} + \mathring{\gamma}_3 \frac{\partial}{\partial x_3}\right)^2,$$

wobei die $\mathring{\gamma}_k$ 4×4reihige Matrizen[1] sind, die der Forderung genügen müssen

$$\mathring{\gamma}_i \mathring{\gamma}_k + \mathring{\gamma}_k \mathring{\gamma}_i = 2\delta_{ik}\,, \tag{1}$$

d. h. gleich der Nullmatrix oder gleich dem doppelten der Einheitsmatrix, je nachdem $i \neq k$ oder $i = k$. Man weiß, daß die $\mathring{\gamma}_k$ durch die Forderung (1) gerade bis auf eine sogenannte Ähnlichkeitstransformation

$$\mathring{\gamma}'_k = S^{-1} \mathring{\gamma}_k S$$

mit einer beliebigen, nichtsingulären 4×4reihigen Transformationsmatrix S bestimmt sind. Diese Freiheit in der Wahl der $\mathring{\gamma}_k$ besteht offenkundig, und man weiß, wie gesagt, daß die Freiheit damit erschöpft ist.

Da man statt vom Wellenoperator auch vom Linienelementquadrat:

$$dx_0^2 + dx_1^2 + dx_2^2 + dx_3^2$$

hätte ausgehen können, liegt es nahe, die Forderungen (1) so aufzufassen, daß die Matrizen $\mathring{\gamma}_k$ neben den anderen Aufgaben, die ihnen nachher bei der Beschreibung des Elektrons zufallen, auch noch die Aufgabe haben, die Weltmetrik zu beschreiben, die vorläufig euklidisch vorausgesetzt war. Soll sie es nicht sein, sondern

$$ds^2 = g_{\mu\nu}\, dx^\mu dx^\nu,$$

[1] Es kommt übrigens für alles Folgende auf die Reihenzahl gar nicht an.

so wird man (1) durch

$$\gamma_i \gamma_k + \gamma_k \gamma_i = 2 g_{ik} \tag{2}$$

zu ersetzen haben (Tetrode). Die γ_k sind Funktionen von Raum und Zeit, d. h. es sind 4×4reihige Matrizen, deren Elemente Funktionen der x_i sind.

Die Gleichungen (2) haben in jedem Punkte P sicherlich Lösungen für die γ_k, wenn man die g_{ik} irgendwie vorgegeben denkt (natürlich so, daß sie einer nichtsingulären Metrik entsprechen). Die Freiheit, welche bei vorgegebenen g_{ik} für die γ_k noch besteht, ist genau dieselbe wie oben für die $\mathring{\gamma}_k$, nämlich: Transformation mit einer beliebigen nichtsingulären Matrix S. Man erkennt die Richtigkeit dieser Behauptungen, indem man sich der Reihe nach folgendes überlegt:

1. Überhaupt lösen lassen sich die Gleichungen (2) immer durch 4 passend gewählte lineare Aggregate eines beliebigen Diracschen Basissystems $\mathring{\gamma}_k$ — der Ansatz führt auf erfüllbare Forderungen für die Koeffizienten.

2. Umgekehrt: Hat man ein System von γ_k, von dem man nur weiß, daß es (2) erfüllt, so kann man 4 Linearaggregate dieser γ_k angeben, die (1) erfüllen, also eine Diracsche Basis bilden. Hat man also etwa zwei Lösungssysteme γ_k und γ_k' von (2), so lassen sie sich durch dieselbe lineare Transformation in je eine Diracsche Basis verwandeln. Diese zwei Diracschen Basen aber gehen sicher durch eine S-Transformation auseinander hervor. Ebendiese verwandelt dann auch γ_k und γ_k' ineinander.

3. Daß jede S-Transformation (2) unangetastet läßt, liegt auf der Hand. — Damit sind die Behauptungen bewiesen.

Ein sehr wesentlicher Unterschied zwischen den $\mathring{\gamma}_k$ und den γ_k ist dieser. Es gibt bekanntlich hermitesche $\mathring{\gamma}_k$-Systeme, hingegen gibt es im allgemeinen keine hermiteschen γ_k-Systeme, auch nicht etwa solche, wo einige γ_k hermitesch, andere schiefhermitesch sind. Das hängt mit den wohlbekannten Reellitätsverhältnissen zusammen, die für die g_{ik} zu fordern sind, nämlich reinimaginär, wenn ein und nur ein Index 0 vorkommt, sonst reell. (Man muß sich erinnern, daß das symmetrische Produkt, der Antikommutator, zweier hermitescher Matrizen stets hermitesch ist.) Wir gehen später auf die Hermitizitätsfragen ausführlich ein und haben dies hier nur erwähnt, um zu zeigen, daß vorläufig nicht der mindeste Grund vorliegt, die in jedem Punkt willkürliche Transformation S etwa auf eine unitäre zu beschränken. Denn da die γ_k ohnedies nicht hermitesch sind, hat man zunächst keinen Anlaß, sich die »Erhaltung der Hermitizität« angelegen sein zu lassen.

Wir leiten jetzt aus (2) ein wichtiges System von Differentialgleichungen für die γ_k ab. Wir denken die g_{ik} vorgegeben und die Gleichungen (2) in jedem Punkt P gelöst, und zwar so, daß diese Lösungen sich zu vier stetigen, differenzierbaren Matrizenfeldern zusammenfügen, was offenbar möglich sein wird.

Gehen wir nun von einem Punkt P zu einem Nachbarpunkt P' über und bilden in diesem Sinne das vollständige Differential der Gleichung (2)

$$\delta\gamma_i\cdot\gamma_k+\gamma_i\cdot\delta\gamma_k+\delta\gamma_k\cdot\gamma_i+\gamma_k\cdot\delta\gamma_i=2\frac{\partial g_{ik}}{\partial x_l}\delta x^l. \tag{3}$$

Beachten wir nun den Satz von Ricci, nach welchem die kovariante Ableitung des Fundamentaltensors g_{ik} identisch verschwindet:

$$g_{ik;l}\equiv\frac{\partial g_{ik}}{\partial x_l}-\Gamma^{\mu}_{kl}g_{i\mu}-\Gamma^{\mu}_{il}g_{\mu k}\equiv 0, \tag{4}$$

so wird die rechte Seite von (3) Folgendem gleich:

$$2(\Gamma^{\mu}_{kl}g_{i\mu}+\Gamma^{\mu}_{il}g_{\mu k})\delta x^l.$$

Diesen Wert kann man der linken Seite von (3) erteilen, wenn man ansetzt:

$$\delta\gamma_i=\Gamma^{\mu}_{il}\gamma_{\mu}\delta x^l \tag{5}$$

und (2) beachtet. D. h. die Matrizen

$$\gamma_i+\delta\gamma_i=\gamma_i+\Gamma^{\mu}_{il}\gamma_{\mu}\delta x^l \tag{6}$$

genügen der Gleichung (2) im Punkte P', wenn die γ_i ihr im Punkte P genügen.

Der Ansatz (5) wäre im allgemeinen widerspruchsvoll, wenn man ihn auf alle Punkte P' in der Umgebung von P anwenden wollte. Denn durch einfache Rechnung kann man sich überzeugen, daß der Ausdruck (5) dann und nur dann ein vollständiges Differential ist, wenn die Krümmung in P verschwindet. Aber nach dem oben Gesagten können und werden sich die γ_i-Werte in P' — wir wollen sie $\gamma_i+\delta'\gamma_i$ nennen — von unserem irgendwie erratenen Lösungsansatz (5) bzw. (6) noch durch eine Ähnlichkeitstransformation unterscheiden, und zwar natürlich durch eine unendlich kleine, wenn die Stetigkeit gewahrt werden soll. Das heißt, es muß eine unendlich kleine Matrix ε geben, derart, daß

$$\gamma_i+\delta'\gamma_i=(1-\varepsilon)(\gamma_i+\delta\gamma_i)(1+\varepsilon)=\gamma_i+\delta\gamma_i+\gamma_i\varepsilon-\varepsilon\gamma_i$$

oder

$$\delta'\gamma_i=\Gamma^{\mu}_{il}\gamma_{\mu}\delta x^l+\gamma_i\varepsilon-\varepsilon\gamma_i. \tag{7}$$

An sich könnte ε für jeden Nachbarpunkt einen anderen, ganz beliebigen Wert haben. Soll aber γ_i einen richtigen Differentialquotienten nach x_l haben, so muß ε für ein Fortschreiten in der Richtung x_l (d. h. für $\delta x_l\neq 0$, alle anderen $=0$) mit δx_l proportional sein. So für jedes l. Soll dann die Änderung von γ_i beim Fortschreiten in beliebiger Richtung sich richtig aus seinen Differentialquotienten berechnen lassen, so muß ε die Summe dieser vier Glieder sein. So kommt man zu dem Ansatz

$$\varepsilon=-\Gamma_l\delta x^l,$$

in welchem die Γ_l vier von Ort und Zeit abhängige Matrizen sind (das Minuszeichen ist natürlich freie Willkür). In (7) eingesetzt, ergibt sich das wichtige System von Differentialgleichungen, das wir oben ankündigten[1]:

$$\frac{\partial \gamma_i}{\partial x_l} = \Gamma_{il}^{\mu} \gamma_\mu + \Gamma_l \gamma_i - \gamma_i \Gamma_l . \tag{8}$$

Wir werden es später so aussprechen: Die kovariante Ableitung des Fundamentalvektors γ_k verschwindet, in voller Analogie zum Satz von Ricci, Gleichung (4). Anderseits hängt die Quellenfreiheit des Viererstroms eng mit diesem Gleichungssystem zusammen. Besonderen Nachdruck möchte ich darauf legen, daß wir es hier rein aus den Anforderungen der Metrik abgeleitet haben, ohne Bezugnahme auf die ψ-Funktion, wobei wir die Transformationsfreiheit der Diracschen Matrizen in Anspruch nehmen mußten. Dadurch traten — und zwar unvermeidlich — die neuen Operatoren Γ_l auf, von denen wir sehen werden, daß sie aufs engste mit dem Viererpotential verknüpft sind (sie bilden aber keinen Vektor!).

Wir untersuchen noch die notwendigen Bedingungen für die Verträglichkeit der Gleichungen (8), nämlich daß die gemischten zweiten Differentialquotienten, auf zwei Arten berechnet, übereinstimmen müssen. Indem man die beim Differenzieren auftretenden ersten Derivierten wieder durch (8) ausdrückt, findet man:

$$\Phi_{kl} \gamma_i - \gamma_i \Phi_{kl} = R_{kli}^{\;\;\;\;\mu} \gamma_\mu . \tag{9}$$

Hier ist $R_{kli}^{\;\;\;\;\mu}$ der gemischte Riemannsche Krümmungstensor in der üblichen Bezeichnungsweise (vgl. z. B. Levi-Civita, Der absolute Differentialkalkül, S. 91; Berlin, bei Springer 1928). Φ_{kl} ist eine Abkürzung, die wir einführen für folgende sechs, in den Indizes k, l antisymmetrische Matrizen:

$$\Phi_{kl} = \frac{\partial \Gamma_l}{\partial x_k} - \frac{\partial \Gamma_k}{\partial x_l} + \Gamma_l \Gamma_k - \Gamma_k \Gamma_l , \tag{10}$$

die, wie sich zeigen wird, in enger Beziehung zum elektromagnetischen Feld stehen. Bei gegebenem γ_i-Feld ist durch (8) jedes Γ_l, durch (9) jedes Φ_{kl} festgelegt bis auf je einen Addenden, der mit allen γ_i vertauschbar, mithin Multiplum der Einheitsmatrix sein muß. Die Φ_{kl} sind aus (9) leicht berechenbar. Man führe neben den γ_i die kontravarianten

$$\gamma^i = g^{ik} \gamma_k \tag{11}$$

ein. Ferner erkläre man

$$s^{\mu\nu} = \tfrac{1}{2}(\gamma^\mu \gamma^\nu - \gamma^\nu \gamma^\mu) . \tag{12}$$

[1] Inhaltlich übereinstimmend mit Fock l. c. Gleichung (24). Die Bedeutung der Zeichen hier und dort ist aber eine etwas andere. Wünscht man zur Deckung zu bringen, so lese man zuerst unseren Abschnitt 5 über Hermitizität!

(Die $s^{\mu\nu}$ entsprechen für μ, $\nu = 1, 2, 3$ in gewisser Weise dem Spin, für $\mu = 0$, $\nu = 1, 2, 3$ in gewisser Weise der Geschwindigkeit. Siehe später.) Wir merken noch an, daß nach (2) und (11)

$$\gamma_i \gamma^k + \gamma^k \gamma_i = 2\,\delta_i^k. \tag{13}$$

Nun findet man leicht

$$\gamma_i s^{\mu\nu} - s^{\mu\nu} \gamma_i = 2(\delta_i^\mu \gamma^\nu - \delta_i^\nu \gamma^\mu). \tag{14}$$

Die $s^{\mu\nu}$ produzieren also beim Vertauschen mit einem γ wieder γ. Das ist gerade, was man zur Auflösung von (9) nach Φ_{kl} braucht. Die rechte Seite von (9) läßt sich ja auch $R_{kl,i\mu}\gamma^\mu$ schreiben, wo $R_{kl,i\mu}$ der symmetrische Riemann-Tensor. Dann bestätigt man mit der V. R. (14), daß

$$\Phi_{kl} = -\tfrac{1}{4} R_{kl,\mu\nu} s^{\mu\nu} + f_{kl} \cdot 1 \tag{15}$$

die allgemeine Lösung von (9) ist[1]. f_{kl} ist das freibleibende Multiplum der Einheit. Die f_{kl} werden (mit i multipliziert) die Rolle des elektromagnetischen Feldes übernehmen. Man erkennt, daß das Auftreten dieser Größen durch den Aufbau der Metrik aus Matrizen zwar sehr nahegelegt wird, daß aber gerade die f_{kl} durch das γ-Feld vorläufig nicht bestimmt sind, sondern von ihm völlig freigelassen werden.

Die $s^{\mu\nu}$ haben als Kommutatoren die Spur Null. Daher ist

$$\text{Spur } \Phi_{kl} = f_{kl} \cdot \text{Spur } 1 = 4 f_{kl}.$$

Anderseits ist nach (10)

$$\text{Spur } \Phi_{kl} = \frac{\partial}{\partial x_k}(\text{Spur } \Gamma_l) - \frac{\partial}{\partial x_l}(\text{Spur } \Gamma_k),$$

weil Differentiation und Spurbildung vertauschbar sind und der Kommutator keinen Beitrag zur Spur liefert. Setzt man etwa

$$\tfrac{1}{4}\,\text{Spur } \Gamma_l = \varphi_l,$$

so wird

$$f_{kl} = \frac{\partial \varphi_l}{\partial x_k} - \frac{\partial \varphi_k}{\partial x_l}. \tag{16}$$

Die Spuren der Γ_l sind das Viererpotential (von einem Faktor i abgesehen).

§ 3. Transformationstheorie, erster Teil.

Nach dem Grundgedanken der allgemeinen Relativitätstheorie soll eine Umbenennung aller Punkte

$$x'_k = x'_k(x_0, x_1, x_2, x_3); \quad k = 0, 1, 2, 3 \tag{17}$$

[1] Inhaltlich wesentlich übereinstimmend mit den indexreichen Beinchengleichungen (46), (48) bei Fock l. c.

die Form der Beschreibung nicht ändern. Dabei soll die Funktion x'_0 nur reinimaginäre, x'_1, x'_2, x'_3 nur reelle Werte annehmen, die Funktionaldeterminante soll positiv bleiben. Wir nennen das eine *Punktsubstitution*. Die g_{ik} transformieren dabei als kovarianter Tensor 2. Stufe.

Solange wir an die γ_i keine andere Anforderung stellen, als daß sie den Gleichungen (2) genügen sollen, läßt sich die Frage, wie sie bei einer Punktsubstitution zu transformieren sind, gar nicht eindeutig beantworten. Denn nach wie vor der Punktsubstitution bleibt eine Ähnlichkeitstransformation mit von Punkt zu Punkt variierender Transformationsmatrix S völlig frei. Wir *können* allerdings bestimmen, daß bei einer reinen Punktsubstitution die γ_i als kovarianter Vektor zu transformieren sind, wodurch (2) jedenfalls erhalten bleibt. Dasselbe ist dann für die Γ_l zu fordern, damit (8) erhalten bleibe. Denn der Kommutator $\Gamma_l \gamma_i - \gamma_i \Gamma_l$ transformiert dann als kovarianter Tensor, welches der restliche Teil der Gleichung, nämlich

$$\frac{\partial \gamma_i}{\partial x_k} - \Gamma^{\mu}_{il} \gamma_{\mu}, \tag{18}$$

ebenfalls tut, wenn γ_i als Vektor substituiert wird; denn (18) ist ja formal die kovariante Ableitung von γ_i. Die Ähnlichkeitstransformationen

$$\gamma'_k = S^{-1} \gamma_k S \tag{19}$$

wären dann als eine Sache für sich zu betrachten, wobei, wie man sich leicht überzeugt, die Γ_l zur Erhaltung von (8) folgendermaßen zu transformieren wären:

$$\Gamma'_l = S^{-1} \Gamma_l S - S^{-1} \frac{\partial S}{\partial x_l}, \tag{20}$$

also *anders* als die γ_k. Hingegen würde man finden, daß nach diesen Festsetzungen die folgenden Aggregate, für die wir das Zeichen ∇_k einführen wollen,

$$\nabla_k = \frac{\partial}{\partial x_k} - \Gamma_k, \tag{21}$$

erstens — selbstverständlich — bei einer reinen Punktsubstitution als kovarianter Vektor substituieren (weil dies ja für die $\frac{\partial}{\partial x_k}$ allein gilt und für die Γ_k festgesetzt wurde) und daß die ∇_k *zweitens* bei einer S-Transformation wegen (20) genau wie die γ_k nach (19) transformieren

$$\nabla'_k = S^{-1} \nabla_k S. \tag{22}$$

Denn die Bedeutung von ∇'_k ist doch

$$\nabla'_k = \frac{\partial}{\partial x_k} - \Gamma'_k = \frac{\partial}{\partial x_k} - S^{-1} \Gamma_k S - S^{-1} \frac{\partial S}{\partial x_k}, \tag{23}$$

und es ist

$$\frac{\partial}{\partial x_k} = \frac{\partial}{\partial x_k} \cdot S^{-1} S = S^{-1} \frac{\partial}{\partial x_k} S - \frac{\partial S^{-1}}{\partial x_k} S = S^{-1} \frac{\partial}{\partial x_k} S + S^{-1} \frac{\partial S}{\partial x_k}, \quad (24)$$

letzteres wegen der Identität:

$$S^{-1} S \equiv 1\,; \quad \frac{\partial S^{-1}}{\partial x_k} S + S^{-1} \frac{\partial S}{\partial x_k} \equiv 0.$$

Durch Eintragen von (24) in (23) bestätigt man (22).

Die durch (10) eingeführten Φ_{kl} würden sich erstens — selbstverständlich — bei einer Punktsubstitution als kovarianter Tensor benehmen, zweitens bei einer S-Transformation analog zu (19)

$$\Phi'_{kl} = S^{-1} \Phi_{kl} S, \quad (25)$$

letzteres wegen (22) und weil sie nach den Definitionen (10) und (21) die Kommutatoren der ∇_k sind:

$$\Phi_{kl} = \nabla_l \nabla_k - \nabla_k \nabla_l\,. \quad (26)$$

Hinzuzufügen wäre noch, daß die Spuren der Φ_{kl}, die f_{kl}, bei der Ähnlichkeitstransformation wegen (25) nicht verändern, diejenigen der Γ_l, die wir φ_l nannten aber wohl, weil für die Γ_l kein mit (19) bzw. (25) analoges Transformationsgesetz gilt, sondern (20).

Wir haben alles dies in der »würde«-Form vorgetragen, weil der getroffenen Festsetzung die eingangs erwähnte Willkür anhaftet: da eine Punktsubstitution im allgemeinen jedenfalls zu einer Abänderung der γ_i zwingt (die alten γ_i werden ja im allgemeinen den Gleichungen (2) nicht mehr genügen!), so steht für die Neuwahl wieder eine ganze Mannigfaltigkeit von γ_i-Feldern zur Verfügung, deren Mitglieder aus einem beliebigen von ihnen durch beliebige, koordinatenabhängige S-Transformationen hervorgehen. Und zunächst ist keines dieser Mitglieder innerlich irgendwie ausgezeichnet, auch nicht das oben ausgewählte.

Es empfiehlt sich nun, mindestens für manche Zwecke, diese Wahlfreiheit weitgehend einzuschränken, indem man sie dazu benutzt, um gewisse, zwar nicht unabweisliche, aber naheliegende Hermitizitätswünsche zu befriedigen, wie man das ja in der speziellrelativistischen Dirac-Theorie ebenfalls zu tun pflegt. Um zu sehen, was sich in dieser Hinsicht erreichen läßt, müssen wir die Eigenwerte der γ_k und ihrer Zweierprodukte näher ins Auge fassen.

§ 4. Eigenwerte und Hermitisierung.

Da nach (2)

$$\gamma_k \gamma_k = g_{kk}\,, \qquad \text{(nicht summieren!)}$$

hat γ_k die Eigenwerte $\pm \sqrt{g_{kk}}$, und zwar jeden zweimal, weil es die Spur Null hat. Man erkennt letzteres, wenn man analog zu (12)

$$s_{\mu\nu} = \tfrac{1}{2}(\gamma_\mu \gamma_\nu - \gamma_\nu \gamma_\mu) \quad (26)$$

setzt. Dann gilt analog zu (14)

$$\gamma^i s_{\mu\nu} - s_{\mu\nu}\gamma^i = 2\,(\delta^i_\mu \gamma_\nu - \delta^i_\nu \gamma_\mu). \tag{27}$$

Jedes γ ist also auf mannigfache Art als Kommutator darstellbar, und ein Kommutator hat immer die Spur Null.

Trotzdem die γ lauter reelle Eigenwerte haben und folglich jedes einzelne von ihnen durch eine S-Transformation hermitesch gemacht werden kann, geht das beispielsweise schon für γ_0 und γ_1 im allgemeinen nicht gleichzeitig, weil ihr symmetrisches Produkt nach (2) gleich $2g_{01}\cdot 1$, also (da g_{01} reinimaginär) schiefhermitesch ist.

Betrachten wir nun weiter die Produkte $\gamma_i\gamma^k$, zunächst für $i \neq k$. Das Quadrat ist [vgl. (13)]:

$$(\gamma_i\gamma^k)^2 = \gamma_i\gamma^k\cdot\gamma_i\gamma^k = -\gamma_i\gamma_i\gamma^k\gamma^k = -g_{ii}g^{kk} \quad \text{(nicht summieren!)}.$$

Die Eigenwerte sind also $\pm i\sqrt{g_{ii}g^{kk}}$, und zwar jeder doppelt, da doch

$$\gamma_i\gamma^k = \tfrac{1}{2}(\gamma_i\gamma^k - \gamma^k\gamma_i)$$

als Kommutator die Spur Null haben muß. — Die Eigenwerte von $\gamma^k\gamma_i$ sind entgegengesetzt gleich, also die nämlichen. — Für $i = k$ dagegen hat man:

$$(\gamma_k\gamma^k)^2 = \gamma_k\gamma^k\gamma_k\gamma^k = \gamma_k(2-\gamma_k\gamma^k)\gamma^k = 2\gamma_k\gamma^k - g_{kk}g^{kk} \quad \text{(nicht s.)}$$

$$(\gamma_k\gamma^k - 1)^2 = 1 - g_{kk}g^{kk}. \quad \text{(nicht s.)}$$

$\gamma_k\gamma^k - 1$ hat also die Eigenwerte $\pm\sqrt{1 - g_{kk}g^{kk}}$, und da es als Kommutator geschrieben werden kann:

$$\gamma_k\gamma^k - 1 = \tfrac{1}{2}(\gamma_k\gamma^k - \gamma^k\gamma_k), \quad \text{(nicht s.)}$$

hat es jeden davon doppelt. $\gamma_k\gamma^k$ hat also die Eigenwerte

$$1 \pm \sqrt{1 - g_{kk}g^{kk}}.$$

und zwar jeden davon doppelt. Für $k = 0$ sind diese Werte reell, da $g_{00}g^{00} \leqq 1$.

Von den 4 Matrizen

$$\gamma_0\gamma^0, \quad \gamma_0\gamma^1, \quad \gamma_0\gamma^2, \quad \gamma_0\gamma^3 \tag{28}$$

hat also die erste nur reelle, die drei anderen reinimaginäre Eigenwerte. Sie besitzen also (von einem Faktor i abgesehen) gerade die Reellitätsverhältnisse eines physikalisch sinnvollen Vierervektors [1]. Es liegt darum nahe zu versuchen,

---

[1] Im euklidischen Fall gehen sie tatsächlich in den Diracschen Stromvektor über (von einem Faktor i abgesehen). Auch die Schwierigkeit, die uns hinderte, die γ_k oder die γ^k selbst zu hermitisieren, nämlich, daß ihre symmetrischen Produkte nicht die erforderlichen Reellitätseigenschaften zeigen, besteht für die Matrizen (28) nicht mehr. Für $i \neq k$ hat man

$$\gamma_0\gamma^i\gamma_0\gamma^k + \gamma_0\gamma^k\gamma_0\gamma^i = \gamma_0(2\delta^i_0 - \gamma_0\gamma^i)\gamma^k + \gamma_0(2\delta^k_0 - \gamma_0\gamma^k)\gamma^i = 2(\delta^i_0\gamma_0\gamma^k + \delta^k_0\gamma_0\gamma^i) - 2g_{00}g^{ik}.$$

Das ist tatsächlich reell, wenn keiner der Indizes i, k gleich Null ist, hingegen wird es für $i = 0$, $k \neq 0$:

$$2\gamma_0\gamma^k - 2g_{00}g^{0k}.$$

Dies hat tatsächlich reinimaginäre Eigenwerte, weil wir es von $\gamma_0\gamma^k$ wissen und g^{0k} reinimaginär ist.

ob diese vier Matrizen gleichzeitig hermitisiert (bzw. schiefhermitisiert) werden können. Es zeigt sich, daß das geht, wobei noch eine Anzahl anderer Matrizen zugleich hermitesch werden. Folgendermaßen.

Wenn der Maßtensor g_{ik} reell und positiv definit ist, so läßt sich den Gleichungen (2) durch hermitesche γ_k genügen, ganz ebenso wie den Gleichungen (1) durch hermitesche $\mathring{\gamma}_k$. Das darf ich wohl ohne Beweis als bekannt ansehen, es handelt sich dabei ja nur um die Projektion eines hermitesch anzusetzenden $\mathring{\gamma}_k$-Systems von einem rechtwinkeligen auf ein schiefes Achsenkreuz, wobei ausschließlich reelle Koeffizienten als Richtungskosinus auftreten. Und da die g_{ik} in diesem Falle reell sind, so fallen dann auch die kontravarianten γ^k hermitesch aus; das heißt, man kann auch den zu (2) analogen kontravarianten Gleichungen

$$\gamma^i \gamma^k + \gamma^k \gamma^i = 2 g^{ik} \tag{29}$$

durch hermitesche γ^k genügen, wenn der Tensor g^{ik} reell und positiv definit ist. Das ist unser Tensor g^{ik} nun allerdings nicht, aber man kann ihn dazu machen, indem man ihn verstümmelt und die »gemischten« raumzeitlichen g^{0k} ($k = 1, 2, 3$) zunächst einmal einfach fortläßt, gleich Null setzt. Sei

$$\alpha^0, \alpha^1, \alpha^2, \alpha^3 \tag{30}$$

ein hermitesches Matrizenquadrupel, welches den Gleichungen (29) mit dem so verstümmelten Maßtensor genügt. Das heißt, es soll

$$\alpha^i \alpha^k + \alpha^k \alpha^i = 2 g^{ik}, \tag{31}$$

wenn keiner oder wenn beide Indizes i, k gleich Null sind, und es soll für $k \neq 0$

$$\alpha^0 \alpha^k + \alpha^k \alpha^0 = 0. \tag{32}$$

Nun setze man

$$\gamma^k = \frac{i}{g^{00}} \alpha^0 \alpha^k \qquad \text{für } k \neq 0 \tag{33}$$

und

$$\gamma^0 = \frac{\alpha^0}{\sqrt{g_{00} g^{00}}} - \frac{1}{g_{00}} (g_{01} \gamma^1 + g_{02} \gamma^2 + g_{03} \gamma^3). \tag{34}$$

Man kann sich durch Ausrechnen überzeugen, daß diese γ^k den unverstümmelten Gleichungen (29) Genüge leisten.

Da nach (32) α^0 mit α^k ($k \neq 0$) schief vertauscht, ist $\alpha^0 \alpha^k$ für $k \neq 0$ schief, daher ist nach (33) $\gamma^1, \gamma^2, \gamma^3$ hermitesch. Ferner berechnet man aus (34)

$$\gamma_0 = g_{0k} \gamma^k = \alpha^0 \sqrt{\frac{g_{00}}{g^{00}}} = \text{hermitesch}. \tag{35}$$

Wir haben also durch unsere Konstruktion mit den kontravarianten γ^1, γ^2, γ^3 zugleich das kovariante γ_0 hermitesch gemacht. — An weiteren Hermitizitäten stellen wir fest: Die kontravarianten reinräumlichen

$$s^{kl} = \tfrac{1}{2}(\gamma^k\gamma^l - \gamma^l\gamma^k) \qquad \text{für } k, l = 1, 2, 3 \tag{36}$$

sind, als Kommutatoren hermitescher Matrizen, schiefhermitesch. Ferner sind für $k \neq 0$ die $\gamma_0\gamma^k$ und ebenso die $\gamma^k\gamma_0$ schief, weil, schon nach (13), γ_0 mit γ^k $(k \neq 0)$ schief vertauscht. Sodann findet man nach (34) und (35), daß $\gamma_0\gamma^0$ und $\gamma^0\gamma_0$ hermitesch sind. Daraus folgt dann weiter sehr leicht durch Herunterziehen des Index, daß für $k \neq 0$ auch $\gamma_0\gamma_k$ und $\gamma_k\gamma_0$ mithin auch

$$s_{0k} = \tfrac{1}{2}(\gamma_0\gamma_k - \gamma_k\gamma_0)$$

schief ausfällt. Man beachte aber ausdrücklich, daß sich über die kovarianten s_{kl} für $k, l \neq 0$, und ebenso über die kontravarianten s^{0k} nichts aussagen läßt! Ebensowenig über γ^0, γ_1, γ_2, γ_3. Wir fassen alle Feststellungen noch einmal zusammen. Nach unserer Konstruktion sind

$$\gamma_0,\ \gamma^1,\ \gamma^2,\ \gamma^3,\ \gamma_0\gamma^0,\ \gamma^0\gamma_0 \ \text{hermitesch}; \tag{37}$$

$$\gamma_0\gamma_k,\ \gamma_k\gamma_0,\ \gamma_0\gamma^k,\ \gamma^k\gamma_0,\ s_{0k},\ s^{kl} \ \text{schief} \qquad (k, l \neq 0).$$

Nun wollen wir uns noch von der Bezugnahme auf die spezielle Matrizenkonstruktion befreien, die nur als Existenzbeweis gedient hat. Es läßt sich leicht einsehen: schon die Forderung, daß vier geeignet ausgewählte von den in (37) angeführten Matrizen die dort festgestellte Eigenschaft haben — beispielsweise die Forderung, daß γ_0, γ^1, γ^2, γ^3 hermitesch ausfallen sollen —, genügt, um bei gegebenen g_{ik} das γ-Feld eindeutig bis auf eine unitäre Transformation festzulegen. Denn mehr Freiheit besteht ja, bei gegebenen g_{ik}, für das γ-Feld überhaupt nicht als: Transformation mit einer beliebigen Matrix. Soll diese Transformation die Matrizen γ_0, γ^1, γ^2, γ^3 hermitesch lassen, aus denen sich jede Matrix, also auch jede hermitesche Matrix, durch Addition und Multiplikation herstellen läßt[1], so muß die Transformation jede hermitesche Matrix hermitesch lassen, d. h. sie muß unitär sein. W. z. b. w.

Wir wollen in Zukunft nur solche γ-Felder — man könnte auch sagen, nur solche Bezugsysteme — zulassen, bei denen die Matrizen γ_0, γ^1, γ^2, γ^3 hermitesch ausfallen. Alles in (37) Festgestellte

[1] Zunächst ist es für die Diracschen $\overset{\circ}{\gamma}_k$ bekannt, daß sich aus ihnen jede Matrix rational herstellen läßt. Danach erschließt man es für die γ_k allein oder für die γ^k allein. Daß in obigem Quadrupel γ^0 fehlt und dafür γ_0 eintritt, schadet nichts, weil doch

$$\gamma_0 = g_{00}\gamma^0 + g_{01}\gamma^1 + g_{02}\gamma^2 + g_{03}\gamma^3,$$

woraus γ^0 berechenbar ist, weil sicher $g_{00} \neq 0$.

gilt dann automatisch. Das »zulässige« Bezugsystem ist durch die Metrik bis auf eine unitäre Transformation bestimmt.

Es ist sehr bequem, durch die neue Forderung die erlaubten S-Transformationen auf unitäre reduziert zu haben, denn diese sind sehr gutmütig und belanglos. Wir brauchen an sie im allgemeinen gar nicht zu denken, können so tun, als wäre das γ-Feld eindeutig durch die Metrik bestimmt. Natürlich erwächst aber jetzt die Aufgabe, wenn man von einem zulässigen γ-Feld ausgeht und eine Punktsubstitution (17) vornimmt, das Transformationsgesetz der γ feiner, nämlich so zu bestimmen, daß man wieder auf ein zulässiges γ-Feld geführt wird. Die im Anfang des Abschnittes 3 gegebene vorläufige Vorschrift: die γ_k sind als kovarianter Vektor zu substituieren — genügt dieser Forderung keineswegs, entspricht ja auch gar nicht dem, wie man in der speziellen Relativitätstheorie vorgeht, wo man die $\mathring{\gamma}_k$ überhaupt nicht substituiert. In der Auffassung des Abschnittes 3 gesprochen könnte man sagen: mit jeder Punktsubstitution muß eine ganz bestimmte (eigentlich: eine bis auf einen unitären Faktor bestimmte!) S-Transformation, die selbst natürlich nicht unitär sein wird, gekoppelt werden, und diese gilt es zu bestimmen. Man kann deshalb füglich von einer ergänzten Punktsubstitution sprechen. Wir werden die Aufgabe im nächsten Abschnitt für unendlich kleine Punktsubstitutionen erledigen.

§ 5. Transformationstheorie, zweiter Teil.

Wir gehen von einem zulässigen γ-Feld aus und gehen zu gestrichenen Variablen über durch die unendlichkleine Punktsubstitution

$$x'_k = x_k + \delta x_k \quad \text{oder} \quad x_k = x'_k - \delta x_k\,, \tag{38}$$

die wir im dargelegten Sinne durch eine unendlichkleine S-Transformation mit

$$S = 1 + \Theta; \quad S^{-1} = 1 - \Theta \tag{39}$$

ergänzen. Wir werden, wie üblich, die Ersetzung der Variablen im Argument nicht zum Ausdruck bringen. Die Gleichungen zwischen gestrichenen und ungestrichenen Operatoren beziehen sich also nicht auf die gleichen, sondern auf entsprechende Argumentwerte, d. h. auf denselben Punkt. — Sei nun zur Abkürzung

$$\frac{\partial\,\delta x_k}{\partial x_l} = a_l^k. \tag{40}$$

Diese Größen sind reinimaginär, wenn ein und nur ein Index gleich Null, sonst reell. Dann wird

$$\begin{aligned} \gamma'_i &= \gamma_i - a_i^l \gamma_l + \gamma_i \Theta - \Theta \gamma_i \\ \gamma'^k &= \gamma^k + a_l^k \gamma^l + \gamma^k \Theta - \Theta \gamma^k\,. \end{aligned} \tag{41}$$

Nimmt man die erste Gleichung für $i = 0$ und multipliziert sie von links in die zweite, so kommt (immer nur in Größen erster Ordnung genau):

$$\gamma_0' \gamma'^k = \gamma_0 \gamma^k - a_0^l \gamma_l \gamma^k + a_l^k \gamma_0 \gamma^l + \gamma_0 \gamma^k \Theta - \Theta \gamma_0 \gamma^k. \tag{42}$$

Wir benutzen das Verfügungsrecht über Θ, um in dieser Gleichung das zweite Glied rechter Hand, das Hermitizitätsschlüsse vereitelt, fortzuschaffen bzw. durch ein anderes zu ersetzen. Das gelingt mit

$$\Theta = -\frac{1}{2 g_{00}} a_0^l \gamma_l \gamma_0. \tag{43}$$

Dann wird nämlich

$$-2\Theta \gamma_0 \gamma^k = a_0^l \gamma_l \gamma^k, \tag{44}$$

und man hat

$$\gamma_0' \gamma'^k = \gamma_0 \gamma^k + a_l^k \gamma_0 \gamma^l + \gamma_0 \gamma^k \Theta + \Theta \gamma_0 \gamma^k.$$

Nun überlege man auf Grund unserer Feststellungen (37), daß nach (43) Θ hermitesch ist. Sein symmetrisches Produkt mit $\gamma_0 \gamma^k$ ist also hermitesch oder schief, je nachdem $\gamma_0 \gamma^k$ es ist. Dasselbe gilt vom zweiten Gliede rechter Hand, es ist schief für $k \neq 0$, hermitesch für $k = 0$. Mithin behalten die $\gamma_0' \gamma'^k$ wirklich dieselbe Hermitizität wie die $\gamma_0 \gamma^k$. Man zeigt ganz ebenso, daß auch γ_0' hermitesch bleibt. Damit ist das γ'-Feld als »zulässig« legitimiert.

Θ ist natürlich nicht eindeutig, aber der in (43) mitgeteilte Wert hat doch die Bedeutung: es ist eindeutig der hermitesche Bestandteil der anzuwendenden unendlichkleinen Matrix. Es könnte ein beliebiger unendlichkleiner schiefer Bestandteil hinzutreten. Man erkennt bei einiger Überlegung, daß er alle Schlüsse ungeändert lassen würde — selbstverständlich, er entspricht ja auch nur einer zusätzlichen unitären Transformation! —

Wir schließen jetzt die scharfe Definition eines Tensoroperators an. *Wenn von einem Operatorensystem*

$$T^{\rho\sigma\cdots}_{\alpha\beta\cdots}$$

bekannt ist oder festgesetzt wird, daß es bei jeder unendlichkleinen ergänzten Punktsubstitution wie ein Tensor von dem durch die Indizes und ihre Stellung angedeuteten Rang transformiert, jedoch unter Hinzutritt des Kommutators

$$T^{\rho\sigma\cdots}_{\alpha\beta\cdots}\Theta - \Theta T^{\rho\sigma\cdots}_{\alpha\beta\cdots},$$

dann wollen wir das Operatorensystem als Tensoroperator von dem betreffenden Range bezeichnen.

Es gilt dann folgender wichtige Satz [1], der durch sehr leichte Verallgemeinerung der obigen Schlüsse gewonnen wird:

---

[1] Die Hermitizitätsbehauptungen haben unmittelbaren Sinn nur dann, wenn $T^{\rho\sigma\cdots}_{\alpha\beta\cdots}$ den Differentiator $\frac{\partial}{\partial x_k}$ nicht enthält, sondern einfach eine 4×4reihige Matrix mit koordinatenabhängigen Elementen ist.

Sei $T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot}$ *ein Tensoroperator und sei bekannt, daß in einem Bezugsystem die Operatoren*

$$\gamma_0 T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot} \tag{46}$$

hermitesch oder schief sind, je nachdem die Null unter den Indizes $\alpha\beta\cdot\cdot\rho\sigma\cdot\cdot$ *in gerader oder ungerader Vielfachheit auftritt; dann bleibt dieser Sachverhalt in jedem Bezugsystem erhalten.* — In diesem Satz darf man selbstverständlich die Worte gerad und ungerad auch vertauschen, d. h. man darf die Null in γ_0 mitzählen oder nicht mitzählen. Was man aber nicht darf, ist, sich um die Hermitizität von $T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot}$ selbst bekümmern, die ist ganz belanglos; auf diejenige von $\gamma_0 T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot}$ kommt es an! —

Man bestätigt leicht, daß das in (21) eingeführte Symbol

$$\nabla_k = \frac{\partial}{\partial x_k} - \Gamma_k \tag{21}$$

ein Vektoroperator ist. Γ_k für sich allein nicht, es transformiert [mit Rücksicht auf (20)] bei einer ergänzten Punktsubstitution offenbar so:

$$\Gamma'_k = \Gamma_k - a^i_k \Gamma_i + \Gamma_k \Theta - \Theta \Gamma_k - \frac{\partial \Theta}{\partial x_k}. \tag{47}$$

Hier ist das letzte Glied überschüssig, widerstreitet der Vektoreigenschaft. Der reine Differentiator $\frac{\partial}{\partial x_k}$ hinwiederum transformiert kovariant im elementaren Sinn, es fehlt der Θ-Kommutator. Fügt man beide zusammen, so gleichen sich diese Übelstände aus, weil je $\frac{\partial \Theta}{\partial x_k}$ als Kommutator von $\frac{\partial}{\partial x_k}$ und Θ aufgefaßt werden kann. — Von »hermitesch« oder »schief« zu sprechen, hat natürlich bei solchen Operatoren wie ∇_k, die Differentiationen enthalten, keinen unmittelbaren Sinn. Wir haben dies ja auch nicht in die Definition eines Tensoroperators aufgenommen.

Hat man zwei Tensoroperatoren, so bestätigt man leicht durch Ausmultiplizieren ihrer Transformationsformeln [ähnlich wie oben beim Übergang von (41) zu (42) geschehen], daß man durch »Aneinanderschreiben«, d. h. durch Matrizenmultiplikation wieder einen Tensor erhält, wenn der links herangeschriebene Operator den Differentialoperator nicht enthält. Sonst nicht, weil er sonst mit den Substitutionskoeffizienten a^k_l nicht vertauschbar ist. (Das ist ja im gewöhnlichen Tensorkalkül auch nicht anders. Obwohl dort $\frac{\partial}{\partial x_k}$ ein Vektor ist, erhält man doch durch gewöhnliche Differentiation von Tensorkomponenten keinen Tensor, sondern durch kovariante Differentiation.) Der Tensorcharakter der Φ_{kl}, die durch (10) oder (26) definiert sind, muß also besonders überlegt werden. Da wir aber schon im § 3 gesehen haben, daß die Φ_{kl} bei reiner Punktsubstitution sich als Tensor im dortgemeinten elementaren Sinn benehmen, bei jeder S-Substitution aber nach (25) transformieren, so bilden sie offenbar auch in dem jetzigen feineren Sinn einen Tensoroperator bei ergänzter Punktsubstitution.

Wir wollen uns jetzt darum kümmern, was unter kovarianter Differentiation eines Tensoroperators zu verstehen sei. Wir beschränken uns dabei auf solche Operatoren, die nicht den Differentialoperator enthalten, also auf 4×4reihige Matrizen, deren Elemente Koordinatenfunktionen sind (das hindert nicht, daß sie die Form von Differentialquotienten haben können; beispielsweise Φ_{kl} ist zugelassen, ∇_k dagegen nicht). Es handelt sich darum, aus einem Tensoroperator $T^{\rho\sigma\cdots}_{\alpha\beta\cdots}$ durch Differentiation nach x_λ und Hinzufügung passender Ergänzungsglieder Gebilde abzuleiten, die bei einer ergänzten Punktsubstitution als ein in $\varrho\sigma\cdots$ kontravarianter, in $\alpha\beta\cdots\lambda$ kovarianter Tensoroperator transformieren.

Wir machen davon Gebrauch, daß eine ergänzte Punktsubstitution äußerlich in eine reine Punktsubstitution und in eine Θ-Transformation zerfällt, bei welch letzterer einfach der Kommutator mit Θ addiert wird; ferner benutzen wir, daß diese beiden unendlichkleinen Transformationen selbstverständlich vertauschbar sind. Betrachten wir nun den kovarianten Differentialquotienten im elementaren Sinn

$$\frac{\partial T^{\rho\sigma\cdots}_{\alpha\beta\cdots}}{\partial x_\lambda} - \Gamma^{\mu}_{\alpha\lambda} T^{\rho\sigma\cdots}_{\mu\beta\cdots} - + \cdots, \tag{48}$$

so wird dieser bei einer reinen Punktsubstitution selbstverständlich als Tensor der Stufe ${}^{\rho\sigma\cdots}_{\alpha\beta\cdots\lambda}$ substituieren. Es würde nur noch nötig sein, zu zeigen, daß er bei Θ-Transformation einfach den Kommutator mit Θ addiert wie $T^{\rho\sigma\cdots}_{\alpha\beta\cdots}$ selber. Das tun nun in dem vorstehenden Ausdruck alle Glieder mit Ausnahme des ersten, in welchem durch die Θ-Transformation das Glied

$$\frac{\partial\left(T^{\rho\sigma\cdots}_{\alpha\beta\cdots}\Theta - \Theta T^{\rho\sigma\cdots}_{\alpha\beta\cdots}\right)}{\partial x_\lambda}$$

hinzugefügt wird an Stelle von

$$\frac{\partial T^{\rho\sigma\cdots}_{\alpha\beta\cdots}}{\partial x_\lambda}\,\Theta - \Theta\,\frac{\partial T^{\rho\sigma\cdots}_{\alpha\beta\cdots}}{\partial x_\lambda}.$$

Es entsteht also überschüssig das Glied

$$T^{\rho\sigma\cdots}_{\alpha\beta\cdots}\frac{\partial\Theta}{\partial x_\lambda} - \frac{\partial\Theta}{\partial x_\lambda}T^{\rho\sigma\cdots}_{\alpha\beta\cdots}. \tag{49}$$

Wir beseitigen es, indem wir in (48) als Ergänzung noch den Kommutator

$$T^{\rho\sigma\cdots}_{\alpha\beta\cdots}\Gamma_\lambda - \Gamma_\lambda T^{\rho\sigma\cdots}_{\alpha\beta\cdots}$$

hinzufügen und so zur endgültigen Definition der kovarianten Differentiation eines Tensoroperators gelangen:

$$T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot;\lambda} = \frac{\partial T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot}}{\partial x_\lambda} - \Gamma^{\mu}_{\alpha\lambda} T^{\rho\sigma\cdot\cdot}_{\mu\beta\cdot\cdot} - + \cdots + T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot} \Gamma_\lambda - \Gamma_\lambda T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot}. \tag{50}$$

Beweis: Das hinzugefügte Glied verhält sich nach (47) so: bei einer reinen Punkttransformation als Tensor der gewünschten Stufe; bei einer Θ-Transformation addiert es erstens seinen Kommutator mit Θ, zweitens beseitigt es den Überschuß (49). Damit ist der Beweis, daß (50) ein Tensor ist, abgeschlossen. Man kann (50) auch in der Form schreiben:

$$T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot;\lambda} = \nabla_\lambda T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot} - T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot} \nabla_\lambda - \Gamma^{\mu}_{\alpha\lambda} T^{\rho\sigma\cdot\cdot}_{\mu\beta\cdot\cdot} - + , \tag{51}$$

welche von der elementaren Formel sich nur dadurch unterscheidet, daß der Differentiator ∇_λ an die Stelle des einfachen $\dfrac{\partial}{\partial x_\lambda}$ getreten ist.

Man erkennt jetzt, daß das wichtige System von Differentialgleichungen (8), auf das wir gleich zu Anfang unserer Überlegungen stießen, nichts weiter aussagt als das Verschwinden der kovarianten Ableitungen des Maßvektors γ_k. In voller Analogie zu dem Satz von Ricci, welcher dasselbe ausspricht für den Maßtensor g_{ik}. Ganz dasselbe gilt übrigens für jeden aus den γ_k durch Multiplikation und Addition mit konstanten Koeffizienten abgeleiteten Tensor, z. B. γ^k, $s_{\mu\nu}$, $s^{\mu\nu}$ usw. Alle diese haben die kovariante Ableitung Null. Das ist eine unmittelbare Konsequenz der Gleichungen (8).

§ 6. Interpretation durch den ψ-Spinor.

Die Beschränkung der γ-Felder auf das, was wir »zulässig« nannten, wird besonders bequem empfunden, wenn zur Interpretation der Operatoren eine vierkomponentige ψ-Funktion, ein sogenannter Spinor, worauf sie einwirken, zugrunde gelegt werden soll. Soll ein Gleichungssystem

$$T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot\cdot}\, \psi = 0 \tag{52}$$

in jedem Bezugsystem richtig bleiben, wenn es in einem gilt, so wird man bestimmen müssen, daß ψ bei einer reinen Punktsubstitution als Invariante, bei einer S-Transformation aber folgendermaßen transformiere:

$$\psi' = S^{-1}\psi. \tag{53}$$

Das erstere ist selbstverständlich. Und bei einer S-Transformation folgt wirklich aus (52) durch Linksmultiplikation mit S^{-1}

$$S^{-1} T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot} S S^{-1} \psi = T'^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot}\, \psi' = 0\,.$$

Bei einer ergänzten unendlichkleinen Punktsubstitution wird man also zu setzen haben

$$\psi' = \psi - \Theta\psi, \tag{54}$$

wo Θ die hermitesche Matrix (43) ist. Da nun ∇_k ein Vektoroperator ist, so folgt unter anderem: wenn die vier Zahlen

$$\nabla_k \psi = \frac{\partial \psi}{\partial x_k} - \Gamma_k \psi \tag{55}$$

in einem Bezugsystem alle vier verschwinden, so tun sie es in jedem. Es ist angebracht, sie als kovariante Ableitungen des Spinors ψ zu bezeichnen.

Aus den Operatoren (q-Zahlen) erhält man die gewöhnlichen Zahlen (c-Zahlen), die physikalisch je nach Geschmack und Ausdrucksweise als Aufenthaltswahrscheinlichkeit, Dichte der Elektrizität, Stromdichte, Übertrittswahrscheinlichkeit u. dgl. interpretiert werden, auf folgende Weise: man übt den betreffenden Operator A auf einen Spinor ψ aus: $A\psi$, und bildet dann das sogenannte hermitesche innere Produkt der beiden Spinoren ψ und $A\psi$, d. h. man multipliziert die erste Komponente der konjugiert komplexen ψ^* mit der ersten von $A\psi$, die zweite von ψ^* mit der zweiten von $A\psi$ usw. und addiert diese 4 Produkte. Wir wollen dafür kurz

$$\psi^* A \psi \tag{56}$$

schreiben[1]. Enthält A den Differentialoperator $\frac{\partial}{\partial x_k}$ nicht, sondern ist einfach nur eine 4×4reihige Matrix mit koordinatenabhängigen Elementen, so kann man auch sagen: man trägt in die aus dieser Matrix gebildete Bilinearform die Komponenten von ψ^* und ψ als Argumente ein.

Nur wenn die Matrix hermitesch (bzw. schief) ist, fällt die c-Zahl (56) immer reell (bzw. reinimaginär) aus, wie es für die Komponenten von c-Tensoren nötig ist, die physikalisch interpretiert werden sollen. Nun haben wir in § 5 gesehen: wenn $T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot}$ ein Tensoroperator ist, so wird bei einer zulässigen Transformation (d. h. bei einer ergänzten Punktsubstitution) keineswegs die Hermitizität seiner Komponenten konserviert, sondern diejenige der $\gamma_0 T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot}$. Es behalten also keineswegs etwa die c-Zahlen $\psi^* T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot}\psi$, sondern die c-Zahlen

$$\mathrm{T}^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot} = \psi^* \gamma_0 T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot} \psi \tag{57}$$

die für einen physikalischen Tensor vom Range ${}^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot}$ erforderlichen Reellitätsverhältnisse bei. Wir wollen jetzt zeigen, daß auch sie es sind, die wirklich wie ein c-Tensor vom Range ${}^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot}$ transformieren und deshalb als physikalische Interpretation des Tensoroperators $T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot}$ zu gelten haben. Man erhält nämlich,

[1] Es kommt in der hier gewählten Schreibweise nicht auf die Reihenfolge an. $A\varphi B\chi$ bedeutet dasselbe wie $B\chi A\varphi$, nämlich immer: erste Komponente von $A\varphi$ mal erster von $B\chi$ plus zweiter von $A\varphi$ mal zweiter von $B\chi$ plus usw.

wenn man die ergänzte Punktsubstitution (38), (40) durchführt, zunächst folgendes:

$$\begin{aligned}\mathsf{T}^{\rho\sigma\cdots\prime}_{\alpha\beta\cdots} &= (\psi^*-\Theta^*\psi^*)(\gamma_0-a^l_0\gamma_l+\gamma_0\Theta-\Theta\gamma_0)(T^{\rho\sigma\cdots}_{\alpha\beta\cdots}-a^l_\alpha T^{\rho\sigma\cdots}_{l\beta\cdots}-+\cdots)(\psi-\Theta\psi)\\ &= \mathsf{T}^{\rho\sigma\cdots}_{\alpha\beta\cdots}-\Theta^*\psi^*\gamma_0 T^{\rho\sigma\cdots}_{\alpha\beta\cdots}\psi-\psi^*\Theta\gamma_0 T^{\rho\sigma\cdots}_{\alpha\beta\cdots}\psi-a^l_0\psi^*\gamma_l T^{\rho\sigma\cdots}_{\alpha\beta\cdots}\psi-\\ &\qquad -a^l_\alpha\mathsf{T}^{\rho\sigma\cdots}_{l\beta\cdots}-+\cdots\end{aligned}\tag{58}$$

(Es haben sich zwei Glieder mit Θ, nämlich das aus $-\Theta\psi$ und das aus $\gamma_0\Theta$ entstehende gegeneinander fortgehoben; Glieder zweiter Ordnung in Θ und a^l_k sind selbstverständlich unterdrückt.) Das zweite, dritte und vierte Glied rechter Hand heben sich fort, denn: das zweite und dritte sind einander gleich, weil Θ^* unter »Transposition« (Vertauschung von Zeilen und Kolonnen) auf den anderen Faktor gewälzt werden darf und dabei, da es hermitesch ist, zu Θ wird. Sodann wird nach (43)

$$-2\psi^*\Theta\gamma_0 T^{\sigma\sigma\cdots}_{\alpha\beta\cdots}\psi = \frac{1}{g_{00}}a^l_0\psi^*\gamma_l\gamma_0\gamma_0 T^{\rho\sigma\cdots}_{\alpha\beta\cdots}\psi = a^l_0\psi^*\gamma_l T^{\rho\sigma\cdots}_{\alpha\beta\cdots}\psi\,,$$

hebt sich also gegen das vierte Glied, wie behauptet. Man erhält also für den c-Tensor (57)

$$\mathsf{T}^{\rho\sigma\cdots\prime}_{\alpha\beta\cdots} = \mathsf{T}^{\rho\sigma\cdots}_{\alpha\beta\cdots}-a^l_\alpha\mathsf{T}^{\rho\sigma\cdots}_{l\beta\cdots}-+\cdots,\tag{59}$$

die gewöhnliche Substitutionsformel, w. z. b. w. — Man beachte ausdrücklich, daß bei diesem Beweis der Operator $T^{\rho\sigma\cdots}_{\alpha\beta\cdots}$ selbst weder gewälzt noch mit einem a^l_k vertauscht zu werden braucht. Der Beweis gilt also auch d a n n, d. h. $\mathsf{T}^{\rho\sigma\cdots}_{\alpha\beta\cdots}$ transformiert auch d a n n als c-Tensor, wenn $T^{\rho\sigma\cdots}_{\alpha\beta\cdots}$ den Differentialoperator $\dfrac{\partial}{\partial x_k}$ enthält. Bloß die Hermitizitätsaussagen haben dann für die lokalen Tensorkomponenten keinen unmittelbaren Sinn.

Für das Folgende ist es bequem, die Formel (55) auszudehnen für den Fall, daß es sich nicht um einen Spinor, sondern um sein Konjugiert-Komplexes handelt. Das Konjugiert-Komplexe von (55) wäre

$$\frac{\partial\psi^*}{\partial x^*_k}-\Gamma^*_k\psi^*,$$

dies würde aber für $k=0$ ($x_0=ict$!) im euklidischen Falle n i c h t in die gewöhnliche Ableitung übergehen, sondern in ihr Negatives, was sehr unbequem wäre. Wir sind darum leider genötigt, für $k=0$ das Vorzeichen zu wechseln und als kovariante Ableitung von ψ^* zu definieren

$$\nabla_k\psi^* = \frac{\partial\psi^*}{\partial x_k}\mp\Gamma^*_k\psi^*\tag{60}$$

(oberes Zeichen für $k=1,2,3$; unteres für $k=0$).

Jetzt wollen wir noch die kovariante Ableitung des c-Tensors (57) untersuchen, die wohl irgendwie mit der in (50) definierten des Tensoroperators zusammenhängen muß. Man erhält zunächst:

$$\begin{aligned}\mathsf{T}^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot;\lambda} &= \frac{\partial \mathsf{T}^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot}}{\partial x_\lambda} - \Gamma^{\mu}_{\lambda\alpha} \mathsf{T}^{\rho\sigma\cdot\cdot}_{\mu\beta\cdot\cdot} - + \cdots = \\ &= \frac{\partial \psi^*}{\partial x_\lambda} \gamma_0 T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot} \psi + \psi^* \frac{\partial \gamma_0}{\partial x_\lambda} T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot} \psi + \psi^* \gamma_0 \frac{\partial T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot}}{\partial x_\lambda} \psi + \psi^* \gamma_0 T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot} \frac{\partial \psi}{\partial x_\lambda} - \\ &\quad - \Gamma^{\mu}_{\lambda\alpha} \psi^* \gamma_0 T^{\rho\sigma}_{\mu\beta} \psi - + \cdots\cdots.\end{aligned}$$

Die vier Ableitungen, die hier vorkommen, ergänze man nun zu kovarianten Ableitungen nach (60), (8), (50), (55), wobei diejenige von γ_0 verschwindet. So erhält man

$$\mathsf{T}^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot;\lambda} = \nabla_\lambda \psi^* \gamma_0 T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot} \psi + \psi^* \gamma_0 T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot;\lambda} \psi + \psi^* \gamma_0 T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot} \nabla_\lambda \psi \tag{61}$$

plus einem Rest, von welchem jetzt zu zeigen sein wird, daß er verschwindet. Dieser Rest ist

$$\begin{aligned}\text{Rest} = {} & \pm \Gamma^*_\lambda \psi^* \gamma_0 T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot} \psi + \\ & + \psi^* [\Gamma^{\mu}_{0\lambda} \gamma_\mu T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot} + (\Gamma_\lambda \gamma_0 - \underline{\gamma_0 \Gamma_\lambda}) T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot} - \underline{\gamma_0 T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot} \Gamma_\lambda + \gamma_0 \Gamma_\lambda T^{\rho\sigma}_{\alpha\beta}}] \psi + \\ & + \underline{\psi^* \gamma_0 T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot} \Gamma_\lambda \psi}.\end{aligned}$$

Die unterstrichenen Glieder tilgen einander. $\pm \Gamma^*_\lambda$ wird als $\pm \Gamma^\dagger_\lambda$ auf den anderen Faktor gewälzt[1]. Es verbleibt

$$\begin{aligned}\text{Rest} &= \psi^* A \, T^{\rho\sigma\cdot\cdot}_{\alpha\beta\cdot\cdot} \psi \text{ mit} \\ A &= \Gamma^{\mu}_{0\lambda} \gamma_\mu + (\Gamma_\lambda \pm \Gamma^\dagger_\lambda) \gamma_0.\end{aligned}$$

Der Beweis wird abgeschlossen sein, wenn wir zeigen können, daß

$$\frac{1}{2 g_{00}} A \gamma_0 \equiv \tfrac{1}{2} (\Gamma_\lambda \pm \Gamma^\dagger_\lambda) + \frac{1}{2 g_{00}} \Gamma^{\mu}_{0\lambda} \gamma_\mu \gamma_0 \tag{62}$$

verschwindet. (Denn daraus folgt $A \equiv 0$, weil γ_0 die nichtverschwindenden Eigenwerte $\pm \sqrt{g_{00}}$ hat. Ist $A = 0$, so verschwindet der »Rest« und Gleichung (61) wird bewiesen sein.)

Der Operator (62) ist nun im Falle des oberen Vorzeichens, welches für $\lambda = 1, 2, 3$ gilt, der hermitesche, im Falle des unteren Vorzeichens der schiefhermitesche Bestandteil von

$$\Gamma_\lambda + \frac{1}{2 g_{00}} \Gamma^{\mu}_{0\lambda} \gamma_\mu \gamma_0. \tag{63}$$

---

[1] Mit dem Kreuz † soll die transponierte und komplex konjugierte Matrix bezeichnet werden, wie fast (leider nur fast) allgemein üblich.

Von diesem Operator läßt sich ohne allzu große Mühe einsehen, daß er, wenn man ihn mit den nach (37) *hermiteschen* Matrizen γ_0, γ^1, γ^2, γ^3 kommutiert, im Falle $\lambda = 1, 2, 3$ *ausnahmslos Hermitesches*, im Falle $\lambda = 0$ *ausnahmslos Schiefhermitesches* liefert. Daher muß sein hermitescher (bzw. für $\lambda = 0$ sein schiefhermitescher) Bestandteil jedenfalls mit γ_0, γ^1, γ^2, γ^3 *vertauschbar*, mithin ein Multiplum der Einheit sein. Anders ausgedrückt, die Bestandteile, um deren Verschwinden es sich handelt, reduzieren sich auf

$$\text{Realteil Spur } \left(\Gamma_\lambda + \frac{1}{2\,g_{00}}\,\Gamma^\mu_{0\lambda}\gamma_\mu\gamma_0\right) \qquad \text{für } \lambda = 1, 2, 3$$

$$\text{Imaginärteil Spur } \left(\Gamma_0 + \frac{1}{2\,g_{00}}\,\Gamma^\mu_{00}\gamma_\mu\gamma_0\right).$$

Nun ist Spur $\gamma_\mu\gamma_0 = 4\,g_{\mu 0}$ und

$$g_{\mu 0}\,\Gamma^\mu_{0\lambda} = \Gamma_{0,\,0\lambda} = \tfrac{1}{2}\,\frac{\partial g_{00}}{\partial x_\lambda}\,; \qquad \text{für } \lambda = 0, 1, 2, 3$$

die Frage dreht sich also darum, ob wirklich

$$\text{Realteil Spur } \Gamma_\lambda = -\frac{\partial \lg g_{00}}{\partial x_\lambda} \qquad \text{für } \lambda = 1, 2, 3; \tag{64}$$

$$\text{Imaginärteil Spur } \Gamma_\lambda = -\frac{\partial \lg g_{00}}{\partial x_0}\,?$$

Und nun zeigt es sich, daß wir zuviel versprochen haben. *Beweisen* können wir nämlich die vorstehenden Gleichungen *nicht*, und zwar aus *dem* Grunde, weil doch die Γ_λ ursprünglich eingeführt und bisher ausschließlich verwendet wurden in *der* Weise, daß *nur ihre Kommutatoren* mit anderen Matrizen eine Rolle spielen, wobei gerade ihre *Spuren* vollkommen leerlaufen. Diese spielen zum ersten Male eine Rolle in der kovarianten Ableitung des *Spinors*, Gleichung (55) und (60), von welcher wir gerade in der zu beweisenden Gleichung (61) zum ersten Male Gebrauch machen. Was wir beweisen können, ist nur, daß es uns frei steht, die betreffenden Spurteile durch (64) zu *definieren*. Und das ist wirklich der Fall. Einmal sicher in *einem* Bezugsystem, weil die rechten Seiten von (64) die erforderliche Reellität besitzen. Sodann läßt sich an Hand von (47) und (43) zeigen, daß die einmal getroffene Verfügung gegen zulässige Transformationen invariant ist — ich unterdrücke den Beweis.

Durch diese Verfügung wird die kovariante Ableitung des Spinors in erwünschter Weise präzisiert. Die Verfügung ist aber noch in anderer Hinsicht wirklich erwünscht. Würden nämlich die in Rede stehenden Spurteile sich nicht als die Ableitungen ein und derselben Funktion ($-\lg g_{00}$) darstellen, so würden sie in den Spuren der Φ_{kl} reinimaginäre elektromagnetische Feldstärken erzeugen. Das wird so vermieden. — Der Realteil von Spur Γ_0 und die Imaginärteile von Spur Γ_λ ($\lambda = 1, 2, 3$), aus denen die reellen Feldstärken sich ableiten, bleiben nach wie vor frei.

Wir müssen jetzt noch einen Blick auf die reinen unitären Transformationen werfen, die für sich allein auch noch zulässig sind neben den ergänzten Punktsubstitutionen. Das einzige, was zu sagen übrigbleibt, ist, daß eine solche unitäre Transformation, die man vornehmen will, selbstverständlich nach der Vorschrift (53) auch an ψ auszuführen ist. Alsdann ist sie völlig belang- und harmlos. Insbesondere sind die Komponenten der c-Tensoren (57) gegen sie völlig unempfindlich; und die Spurteile, über die in (64) verfügt wurde, ebenfalls.

Die wesentlichen Ergebnisse dieses Abschnittes sind:

1. Die Bestimmung des Transformationsgesetzes (54) und der kovarianten Ableitung (55) für den Spinor.

2. Die Zuordnung der c-Tensorkomponenten zum Tensoroperator nach (57) und der Nachweis, daß sie wirklich als gewöhnliche Tensorkomponenten vom gleichen Range transformieren.

3. Die Aufstellung einer relativ einfachen Formel (61) zur Berechnung der kovarianten Ableitung eines c-Tensors; einer Formel, die hauptsächlich um deswillen von Interesse ist, weil sie zu ihrer Gültigkeit erfordert die an sich willkommene

4. Normierung desjenigen Spurbestandteils von Γ_λ, der ohne Normierung zum Auftreten reinimaginärer elektromagnetischer Feldstärken Anlaß geben würde.

§ 7. Die Diracsche Gleichung.

Der Operator $\gamma^k \nabla_k$ ist eine Invariante, die man passend als »Betrag des Gradienten« bezeichnen kann. Die verallgemeinerte Diracsche Gleichung fordert[1]

$$\gamma^k \nabla_k \psi = \mu \psi, \tag{65}$$

[1] Man könnte allerdings versucht sein, zu »symmetrisieren« und als linke Seite von (65) zu nehmen

$$\tfrac{1}{2}(\gamma^k \nabla_k + \nabla_k \gamma^k). \tag{66}$$

Dieser Ausdruck läßt sich aber umformen. Das Verschwinden der kovarianten Ableitung von γ^k sagt aus:

$$\nabla_l \gamma^k - \gamma^k \nabla_l = -\Gamma_{l\mu}^{k} \gamma^\mu.$$

Durch Verjüngen entsteht

wo μ eine universelle Konstante

$$\mu = \frac{2\pi m c}{h}\,.$$

Der nach der Zuordnung (57) zu γ^k gehörige c-Vektor heiße iS^k, also

$$iS^k = \psi^* \gamma_0 \gamma^k \psi. \tag{69}$$

Da die kovariante Ableitung des Operators γ^k verschwindet, reduziert sich diejenige von S^k nach (61) auf

$$iS^k{}_{;\lambda} = \nabla_\lambda \psi^* \gamma_0 \gamma^k \psi + \psi^* \gamma_0 \gamma^k \nabla_k \psi\,.$$

Bildet man durch Verjüngung die kovariante Divergenz:

$$iS^\lambda{}_{;\lambda} = \nabla_\lambda \psi^* \gamma_0 \gamma^\lambda \psi + \psi^* \gamma_0 \gamma^\lambda \nabla_\lambda \psi,$$

so ist der erste Summand das Negativgenommene, Konjugiert-Komplexe des zweiten[1], dieser aber ist nach (65)

$$\mu \psi^* \gamma_0 \psi,$$

also reell, weil γ_0 hermitesch. Also ist

$$S^\lambda{}_{;\lambda} = 0\,. \tag{70}$$

So folgt die Quellenfreiheit des Viererstroms, der nach unserer Zuordnung (57) als c-Vektor zum kontravarianten Maßvektor gehört, aus der Dirac-Gleichung und den fundamentalen Gleichungen (8) (vgl. Fock l. c. S. 267).

Wir wollen jetzt die Dirac-Gleichung quadrieren, um das Ergebnis mit dem aus der speziellen Theorie vertrauten zu vergleichen (der Kürze halber werde ψ unterdrückt):

$$\gamma^k \nabla_k \gamma^l \nabla_l = \mu^2\,. \tag{71}$$

$$\nabla_k \gamma^k - \gamma^k \nabla_k = -\Gamma_{k\mu}^{\,k} \gamma^\mu = -\frac{\partial \lg \sqrt{g}}{\partial x_\mu}\gamma^\mu\,. \tag{67}$$

Folglich ist

$$\tfrac{1}{2}(\gamma^k \nabla_k + \nabla_k \gamma^k) = \gamma^k \nabla_k - \tfrac{1}{2}\frac{\partial \lg \sqrt{g}}{\partial x_k}\gamma^k = g^{\frac{1}{4}} g^k \nabla_k g^{-\frac{1}{4}}\,. \tag{68}$$

Dies ist kein invarianter Operator, was uns nicht zu wundern braucht. $\nabla_k \gamma^k$ ist nämlich keiner und hat auch keine Verpflichtung, es zu sein. Denn wir haben schon oben hervorgehoben, daß das Produkt zweier Tensoroperatoren nur dann sicher Tensoreigenschaft hat, wenn der linke Faktor den Differentiator nicht enthält. Im übrigen würde die Verwendung des Ansatzes (66) doch wieder auf dasselbe hinauslaufen, man müßte bloß $g^{-\frac{1}{4}}\psi$ an die Stelle von ψ treten lassen, das heißt, man müßte $g^{-\frac{1}{4}}\psi$ als Spinor transformieren. Wir bleiben deshalb beim Ansatz (65).

[1] Das hermitesche $\gamma_0 \gamma^0$ wälzt sich als $(\gamma_0 \gamma^0)^*$ auf den ersten Faktor, die schiefen $\gamma_0 \gamma^{\lambda \neq 0}$ als $-(\gamma_0 \gamma^\lambda)^*$. Dafür enthält aber ∇_0 einen Vorzeichenwechsel, $\nabla_{\lambda \neq 0}$ nicht. Vergleiche das oben zu Gleichung (60) im Text Bemerkte sowie auch die Anmerkung zu Gleichung (56).

Man vertausche die ersten beiden Faktoren mittels Gleichung (67) (in der Anmerkung) und verwende, daß nach (2) und (12)

$$\gamma^k \gamma^l = g^{kl} + s^{kl}. \tag{72}$$

So kommt

$$\nabla_k (g^{kl} + s^{kl}) \nabla_l + \frac{\partial \lg \sqrt{g}}{\partial x_\mu} \gamma^\mu \gamma^l \nabla_l = \mu^2.$$

Aus dem Verschwinden der kovarianten Ableitung von s^{kl} folgt

$$\nabla_k s^{kl} - s^{kl} \nabla_k = - \frac{\partial \lg \sqrt{g}}{\partial x_\mu} s^{\mu l}.$$

So kommt [mit nochmaliger Verwendung von (72)]:

$$\nabla_k g^{kl} \nabla_l + s^{kl} \nabla_k \nabla_l + \frac{\partial \lg \sqrt{g}}{\partial x_\mu} g^{\mu l} \nabla_l = \mu^2.$$

Das zweite Glied ist nach (26) und wegen der Antisymmetrie der s^{kl} gleich $-\frac{1}{2} s^{kl} \Phi_{kl}$. Das erste und dritte (wo man μ durch k ersetze) vereinigen sich zur verallgemeinerten Laplaceschen Operation; man erhält also schließlich:

$$\frac{1}{\sqrt{g}} \nabla_k \sqrt{g}\, g^{kl} \nabla_l - \frac{1}{2} s^{kl} \Phi_{kl} = \mu^2. \tag{73}$$

Es ist interessant, hier für Φ_{kl} den viel früher gefundenen Ausdruck (15) einzusetzen. Dabei tritt die Invariante

$$\frac{1}{8} R_{kl,\mu\nu} s^{kl} s^{\mu\nu}$$

auf. Wegen der Symmetrie des kovarianten Riemannschen Krümmungstensors im ersten und zweiten Indexpaar ist das auch gleich

$$\frac{1}{16} R_{kl,\mu\nu} (s^{kl} s^{\mu\nu} + s^{\mu\nu} s^{kl}).$$

Wenn man nun — was ich hier nicht in extenso durchführen will — die symmetrischen Produkte der s^{kl} wirklich ausrechnet, sodann von der bekannten zyklischen Symmetrie

$$R_{kl,\mu\nu} + R_{l\mu,k\nu} + R_{\mu k,l\nu} = 0$$

Gebrauch macht, erhält man schließlich

$$\frac{1}{8} R_{kl,\mu\nu} s^{kl} s^{\mu\nu} = -\frac{1}{4} g^{k\mu} g^{l\nu} R_{kl,\mu\nu} = -\frac{R}{4},$$

wo R die invariante Krümmung. Mithin ergibt das Einsetzen von Φ_k nach (15) in (73) folgendes:

$$\frac{1}{\sqrt{g}}\nabla_k \sqrt{g}\, g^{kl}\nabla_l - \frac{R}{4} - \frac{1}{2} f_{kl} s^{kl} = \mu^2. \tag{74}$$

In dem dritten Gliede linker Hand erkennt man die wohlvertraute Einwirkung der Feldstärke auf den Spintensor, und zwar ist in f_{kl} bereits der reine Spuranteil von Φ_{kl} abgelöst, der also wohl im eigentlichen Sinne als Feldstärke zu bezeichnen ist und, wie öfters erwähnt, durch die Metrik noch völlig freigelassen wird.

Das zweite Glied scheint mir von erheblichem theoretischen Interesse. Es ist freilich um viele, viele Zehnerpotenzen zu klein, um etwa das Glied rechter Hand ersetzen zu können. Denn μ ist die reziproke Compton-Wellenlänge, ungefähr $10^{11}\,\mathrm{cm}^{-1}$. Immerhin scheint es bedeutungsvoll, daß in der verallgemeinerten Theorie überhaupt ein mit dem rätselhaften Massenglied gleichartiges ganz von selber angetroffen wird[1].

[1] S. a. O. Veblen und B. Hoffmann, Phys. Rev. **36**, 821, 1930.

Ausgegeben am 30. April.

Chapter 14
Leopold Infeld and Bartel van der Waerden (1933): Die Wellengleichung des Elektrons in der allgemeinen Relativitätstheorie

Leopold Infeld and Bartel van der Waerden (1933). Die Wellengleichung des Elektrons in der allgemeinen Relativitätstheorie. *Sitzungsberichte der preussischen Akademie der Wissenschaften, physikalisch-mathematische Klasse*, 380–401 (errata on p. 474).

Die Wellengleichung des Elektrons in der allgemeinen Relativitätstheorie.

Von Priv.-Doz. Dr. L. Infeld
in Lemberg

und Prof. Dr. B. L. van der Waerden
in Leipzig.

(Vorgelegt von Hrn. Schrödinger am 12. Januar 1933 [s. oben S. 2].)

§ 1. Einleitung.

Man hat früher gemeint, daß der Einbau der Diracschen Theorie des Elektrons in die allgemeine Relativitätstheorie nur auf Grund eines Fernparallelismus möglich wäre[1]. Diese Meinung hat V. Fock[2] widerlegt, indem er in bezug auf ein beliebiges metrisches Feld in jedem Punkt willkürlich ein orthogonales n-Bein einführt und die Unabhängigkeit der aufgestellten Gleichungen von der Wahl des n-Beins dartut. Die etwas ungeläufige n-Beinrechnung sucht Schrödinger[3] zu vermeiden, indem er die Diracschen vierreihigen Matrices nach dem Beispiel von Tetrode verallgemeinert und von Punkt zu Punkt variabel macht. Jedoch scheint uns der Schrödingersche Rechenapparat unnötig kompliziert, was sich auch darin äußert, daß bei Schrödinger die Transformationsweise der betrachteten Größen nicht einmal explizit angegeben wird, sondern man sich mit der Angabe der infinitesimalen Transformationen begnügen muß.

In der vorliegenden Arbeit wird ein Rechenapparat entwickelt, der unter Vermeidung der erwähnten Mängel dasselbe leistet wie die Theorien von Fock und Schrödinger. Statt vierreihiger werden nur zweireihige Matrices eingeführt, zu deren Darstellung die gewöhnlichen Regeln der Tensorrechnung ausreichen. Neben den Weltvektoren treten zweikomponentige »Spinvektoren« auf, welche (in Gegensatz zur früheren »speziellen« Spinoranalyse[4]) unabhängig von den Weltvektoren transformiert werden, wodurch die formalen Regeln zur Bildung invarianter Gleichungen bekanntlich sehr einfach werden. Als Bindeglied zwischen Weltvektoren und Spinoren dienen gewisse gemischte Größen $\sigma^{k\dot{\lambda}\mu}$ (verallgemeinerte Paulische Matrices), welche

[1] Vgl. etwa E. Wigner, Z. f. Phys. 53, 592, 1929.
[2] V. Fock, Z. f. Phys. 57, 261, 1929. Vgl. auch H. Weyl, Z. f. Phys. 56, 330, 1929.
[3] E. Schrödinger, S.-B. preuß. Akad. Wiss. 1932, 105.
[4] Siehe B. L. van der Waerden, Göttinger Nachrichten 1929, 100.

die Rolle der Fockschen n-Beinkomponenten übernehmen und ebenso wie diese das metrische Feld g_{kl} mitbestimmen. Nach Einführung einer kovarianten Differentiation für Spinoren lassen sich die Diracschen Gleichungen unmittelbar hinschreiben. In dem Ausdruck für die kovariante Differentiation tritt ein willkürlicher Weltvektor Φ_k auf, der mit dem elektromagnetischen Potential identifiziert wird. Das elektromagnetische Feld wird dadurch ähnlich wie in den bekannten einheitlichen Feldtheorien von Gravitation und Elektrizität geometrisiert.

Unsere Arbeit weist viele Berührungspunkte mit einer Arbeit von J. A. Schouten[1] auf. Schouten kommt am Ende nahezu zum gleichen Formalismus, der in dieser Arbeit entwickelt wird; nur benutzt er zur Einführung dieses Formalismus unnötigerweise n-Beinkomponenten und Sätze über Sedenionen, während der Formalismus nachher noch durch Hilfsvariablen und Pseudogrößen belastet wird. Die Einführung von »Spinordichten« haben wir von Schouten übernommen.

Es kommen in der vorliegenden Arbeit zwei Formalismen nebeneinander zur Darstellung: der einfachere »γ-Formalismus«, in welchem ein alternierender Spinor $\gamma_{\lambda\mu}$ zugrunde gelegt wird, und der etwas kompliziertere »ε-Formalismus«, welcher in Anlehnung an Schouten Spinordichten benutzt. In beiden Formalismen gilt das Prinzip der Eichinvarianz der Potentiale. Daß die Einführung von »Pseudogrößen« zum Zwecke der Eichinvarianz hier entbehrlich wird, liegt daran, daß wir (in Gegensatz zu Schouten) im Spinraum auch Transformationen mit komplexer Determinante zulassen: diese ersetzen die Weyl-Schoutenschen Umeichungen. Die Kenntnis der in der Einleitung zitierten Arbeiten ist für die Lektüre des Folgenden nicht erforderlich.

§ 2. Der metrische Raum und der Spinraum.

Wir legen zunächst die gewohnten Begriffe der allgemeinen Relativitätstheorie zugrunde: vierdimensionaler Riemannscher Raum, metrischer Tensor mit den Komponenten g_{kl} (g_{11} bis $g_{33} < 0$, $g_{44} > 0$), kovariante und kontravariante Weltvektoren und Welttensoren.

Jedem Punkt des metrischen Raumes mit den reellen Koordinaten x^1, x^2, x^3, x^4 soll nun weiter ein komplexer zweidimensionaler Spinraum zugeordnet werden, dessen Vektoren und Tensoren mit dem Sammelnamen Spinoren bezeichnet werden. Wir verabreden: Die Indices der Welttensoren werden mit lateinischen, diejenigen der Spinoren mit griechischen Buchstaben bezeichnet. Bei zweimal auftretenden Indices ist stets zu summieren, und zwar bei den lateinischen von 1 bis 4, bei griechischen von 1 bis 2.

[1] J. A. Schouten, Journal of Math. and Phys. 10, 239, 1931.

Es seien $a^\lambda (\lambda = 1, 2)$ die Komponenten eines kontravarianten Spinvektors. Sie transformieren sich bei Koordinatentransformationen im Spinraume in folgender Weise:

$$a'^\lambda = \Lambda^\lambda_\rho a^\rho. \tag{1a}$$

Sowohl a^λ als auch Λ^λ_ρ sind im allgemeinen komplexe Funktionen der als Parameter auftretenden Weltpunkte. Die einzige Bedingung (außer der Differenzierbarkeit), denen die Transformationskoeffizienten Λ^λ_ρ zu genügen haben, ist die folgende: Die Determinante $|\Lambda^\lambda_\rho|$ soll überall von Null verschieden sein:

$$\Delta = \mathrm{Det}|\Lambda^\lambda_\rho| \neq 0. \tag{2}$$

Die Transformationen des vierdimensionalen metrischen Raumes und des Spinraumes sind als voneinander völlig unabhängig zu betrachten.

Der zu a^λ konjugiert komplexe Spinvektor soll durch $a^{\dot\lambda}$ bezeichnet werden. Seine Transformationsgleichungen sind:

$$a'^{\dot\lambda} = \overline{\Lambda^\lambda_\rho}\, a^{\dot\rho}. \tag{1b}$$

Die Spinoren

$$\beta^{\dot\lambda\mu}, \quad \beta^{\lambda\dot\mu}, \quad \beta^{\dot\lambda\mu\nu} \quad \text{usw.}$$

transformieren sich entsprechend wie

$$a^{\dot\lambda} a^\mu, \quad a^\lambda a^{\dot\mu}, \quad a^{\dot\lambda} a^\mu a^\nu \quad \text{usw.}$$

Die Spinoren $\beta^{\dot\lambda\mu\dot\nu}$ und $\beta^{\lambda\dot\mu\nu}$ sind zueinander konjugiert komplex, da alle nicht punktierten Indices durch punktierte ersetzt wurden und umgekehrt. Es ist also:

$$\beta^{\dot\lambda\mu\dot\nu} = \overline{\beta^{\lambda\dot\mu\nu}}. \tag{3}$$

Es war bis jetzt in der Spinoranalyse üblich, sich auf Transformationen mit konstanten Koeffizienten und von der Determinante Eins im Spinraum zu beschränken. Dabei haben zwei Spinvektoren a^λ, β^λ die absolute Invariante

$$\begin{gathered} \varepsilon_{\lambda\mu} a^\lambda \beta^\mu = a^1 \beta^2 - a^2 \beta^1 \\ (\varepsilon_{12} = -\varepsilon_{21} = 1; \quad \varepsilon_{11} = \varepsilon_{22} = 0). \end{gathered} \tag{4}$$

Wenn man beliebige lineare Transfomationen zuläßt, ist dieser Ausdruck nicht mehr absolut invariant. Da man aber die Bildung $\varepsilon_{\lambda\mu} a^\mu$, die das Herauf- und Herunterziehen der Indices gestattet, in der Spinorenrechnung nicht gut entbehren kann, so wird eine schiefsymmetrische Größe $\gamma_{\lambda\mu} = -\gamma_{\mu\lambda}$ eingeführt, welche die Funktion der $\varepsilon_{\lambda\mu}$ übernimmt und im Spinraum eine ähnliche Rolle spielt wie der metrische Fundamentaltensor g_{ik} im x-Raum[1].

[1] Vgl. L. Infeld, Phys. ZS. 33, 475, 1932.

Wegen $\gamma_{\lambda\mu} = -\gamma_{\mu\lambda}$ ist $\gamma_{12} = -\gamma_{21}$ die einzige von Null verschiedene Komponente von $\gamma_{\lambda\mu}$; diese kann aber eine ganz beliebige von Null verschiedene Funktion des Weltpunktes sein. Die konjugiert-komplexe Größe sei $\gamma_{\dot\lambda\dot\mu}$. In allen physikalisch wichtigen Formeln tritt nur das Produkt $\gamma_{\lambda\mu}\gamma_{\dot\varrho\dot\sigma}$ auf. Dieses Produkt definiert eine invariante Volummessung im Spinraum[1].

Die inverse Matrix zur Matrix $\gamma_{\lambda\mu}$ ist $\gamma^{\lambda\mu}$, wo

$$\gamma^{12} = -\gamma^{21} = \frac{1}{\gamma_{12}}, \quad \gamma^{11} = \gamma^{22} = 0 \tag{5}$$

ist. Den Spinor $\gamma_{\lambda\mu}$ charakterisiert eine komplexe Zahlengröße. Es sei daher:

$$\gamma_{12}\gamma_{\dot1\dot2} = \gamma, \qquad \gamma_{12} = \sqrt{\gamma}\, e^{i\vartheta}, \qquad \gamma_{\dot1\dot2} = \sqrt{\gamma}\, e^{-i\vartheta}.$$

Mit Hilfe der Größen $\gamma^{\lambda\mu}$ und $\gamma_{\lambda\mu}$ kann man nun den Übergang von den kovarianten zu den kontravarianten Komponenten eines Vektors oder Tensors definieren:

$$\left.\begin{aligned} a_\mu &= a^\varrho \gamma_{\varrho\mu}, & a^\mu &= \gamma^{\mu\varrho} a_\varrho \\ a_{\dot\mu} &= a^{\dot\varrho} \gamma_{\dot\varrho\dot\mu}, & a^{\dot\mu} &= \gamma^{\dot\mu\dot\varrho} a_{\dot\varrho} \\ a_{\dot\lambda\mu} &= a^{\dot\varrho\sigma} \gamma_{\dot\varrho\dot\lambda}\gamma_{\sigma\mu}, \text{ usw.} \end{aligned}\right\} \tag{6}$$

Das skalare Produkt

$$-a_\varrho \beta^\varrho = a^\varrho \beta_\varrho = \gamma^{\varrho\sigma} a_\sigma \beta_\varrho = \gamma_{\varrho\sigma} a^\sigma \beta^\varrho$$

ist invariant; speziell ist

$$a_\lambda a^\lambda = 0.$$

Die Annahme der Größe $\gamma_{\lambda\mu}$ ist zwar der formal einfachste Weg, das Herunterziehen der Indices zu ermöglichen, aber nicht der einzig mögliche. Eine zweite Möglichkeit bietet der Begriff der Spinordichte, den wir jetzt erklären wollen.

Eine (skalare) Spindichte vom Gewicht $\mathfrak{k}$ ist eine solche Zahlengröße, welche sich bei Raumtransformationen wie ein Skalar transformiert, aber bei Spintransformationen von der Determinante Δ mit $\Delta^{-\mathfrak{k}}$ multipliziert wird, wobei $\mathfrak{k}$ eine ganze Zahl ist. Eine Spinordichte (Spinvektordichte, Spintensordichte) vom Gewicht $\mathfrak{k}$ ist eine Größe, die sich transformiert wie ein Produkt aus einem Spinor (Spinvektor, Spintensor) und einer Spindichte. Z. B. ist die in (4) eingeführte Größe $\varepsilon_{\lambda\mu}$ eine Spinordichte vom Gewicht -1.

Benutzt man nun den Begriff der Spinordichte, so kann man auch $\varepsilon_{\lambda\mu}$ statt $\gamma_{\lambda\mu}$ zum Hinunterziehen der Indices benutzen; nur ist dann

$$a_\mu = a^\varrho \varepsilon_{\varrho\mu}$$

kein Vektor, sondern eine Vektordichte vom Gewicht -1, wenn a^ϱ ein Spinvektor ist. Ebenso ist natürlich die inverse Größe $\varepsilon^{\lambda\mu}$ eine Spinordichte vom Gewicht $+1$.

Diejenigen Größen, die bei Spintransformationen den Faktor $\overline{\Delta}^{-\mathfrak{k}}$ annehmen, werden wir als Spindichten vom Gewicht $\dot{\mathfrak{k}}$ ansprechen. So ist die zu $\varepsilon_{\lambda\mu}$ konjugiert-komplexe Größe $\varepsilon_{\dot\lambda\dot\mu}$ (die übrigens in jedem Koordinatensystem dieselben Komponenten ± 1 und 0 hat, wie $\varepsilon_{\lambda\mu}$ und $\varepsilon^{\lambda\mu}$) eine Spinordichte vom Gewicht $-\dot{1}$. Schließlich brauchen wir noch solche Größen, welche bei Spin-

[1] Das Volumelement ist nämlich durch $\gamma_{12}\gamma_{\dot1\dot2}\dfrac{\partial(\mu^1, \mu^2, \mu^{\dot1} \mu^{\dot2})}{\partial(q, r, s, t)}\, dq\, dr\, ds\, dt$ gegeben, wo q, r, s, t reelle Parameter sind, von denen die komplexen Spinvariablen μ^1, μ^2 irgendwie abhängen.

transformationen den Faktor $|\Delta|^{-\mathfrak{k}}$ annehmen (hier kann $\mathfrak{k}$ auch gebrochen sein); diese nennen wir **Spindichten** (bzw. Spinordichten) **vom Absolutgewicht** $\mathfrak{k}$. So ist $\varepsilon_{\lambda\mu}\,\varepsilon_{\dot\rho\dot\sigma}$ eine Spinordichte vom Absolutgewicht -2 und $\varepsilon^{\lambda\mu}\,\varepsilon^{\dot\rho\dot\sigma}$ eine vom Absolutgewicht $+2$.

Wie man sieht, ist die Einführung der Dichten eine formale Belastung; sie hat aber den Vorteil, daß die Einführung des Spinors $\gamma^{\lambda\mu}$, dessen physikalische Realität (vgl. später in § 4) nicht feststeht, vermieden wird. Wir werden im folgenden die beiden besprochenen Möglichkeiten als γ-**Formalismus** und ε-**Formalismus** unterscheiden. Welche von diesen Formalismen den Vorzug verdient, das können wir der weiteren Entwicklung der Theorie überlassen.

Es sei noch bemerkt, daß jede algebraische Beziehung zwischen Spinoren, die im γ-Formalismus formuliert wurde, in den ε-Formalismus übertragen werden kann. Dividiert man und multipliziert zugleich z. B. $\gamma_{\lambda\mu}$ durch $\gamma_{12} = e^{i\vartheta}\sqrt{\gamma}$, dann kann man $\frac{\gamma_{\mu\lambda}}{\sqrt{\gamma}}\,e^{-i\vartheta}$ durch $\varepsilon_{\lambda\mu}$ und die Spinoren durch entsprechende Spinordichten ersetzen.

Bei den Raumtransformationen transformieren sich nur die Welttensoren, bei den Transformationen des Spinraumes nur die Spinoren. Es können aber außer den Weltvektoren und Spintensoren noch **gemischte Größen** eingeführt werden. Eine gemischte Größe von der Art

$$\sigma^k_{\dot\lambda\mu}$$

verhält sich bei den Raumtransformationen wie ein kontravarianter Weltvektor, bei den Transformationen des Spinraumes wie ein Spintensor. Ganz allgemein werden die gemischten Größen in den lateinischen Indices wie Welttensoren, in den griechischen Indices wie Spinoren transformiert.

§ 3. Weltvektoren und Spintensoren.

Die Möglichkeit, die Spinoren überhaupt zu räumlichen Größen in Beziehung zu setzen, beruht bekanntlich[1] darauf, daß man jedem Hermitesch symmetrischen Spintensor

$$a_{\dot\lambda\mu} = a_{\dot\mu\lambda} \tag{7}$$

(mit zwei reellen Komponenten $a_{\dot 1 1}$, $a_{\dot 2 2}$ und zwei konjugiert-komplexen $a_{\dot 1 2}$ und $a_{\dot 2 1}$) eineindeutig einen reellen Weltvektor a^k zuordnen kann, dessen 4 Komponenten lineare Funktionen der $a_{\dot\lambda\mu}$ sind:

$$a^k = \sigma^{k\dot\lambda\mu}\,a_{\dot\lambda\mu}\,. \tag{8}$$

Und zwar wird diese Zuordnung so eingerichtet, daß den linearen Transformationen der $a_{\dot\lambda\mu}$, welche diese bei einer Koordinatentransformation mit konstanten Koeffizienten und von der Determinante Eins im Spinraum erleiden, die Lorentztransformationen der zugeordneten Vektoren a^k entsprechen.

Im Fall der speziellen Relativitätstheorie, wo die (die Lorentzgruppe bestimmenden) g_{kl} numerische Konstanten sind, konnte man für die $\sigma^{k\dot\lambda\mu}$ auch

---

[1] Vgl. B. L. van der Waerden, Spinoranalyse, Nachr. Ges. Wiss. Göttingen 1929, S. 100, oder Die Gruppentheoretische Methode in der Quantenmechanik, Berlin 1932, S. 82, § 20.

numerische Konstanten wählen. Im Fall der allgemeinen Relativitätstheorie geht das natürlich nicht mehr, sondern es werden die $\sigma^{k\dot{\lambda}\mu}$ als beliebige Ortsfunktionen angenommen, welche nur der Bedingung genügen müssen, eine eineindeutige Beziehung zwischen den reellen Vektoren a^k und den symmetrischen Tensoren $a_{\dot{\lambda}\mu}$ zu vermitteln.

Damit in (8) die a^k stets reell ausfallen, muß man die gemischte Größe $\sigma^{k\dot{\lambda}\mu}$ auch Hermitesch symmetrisch wählen:

$$\sigma^{k\dot{\lambda}\mu} = \overline{\sigma^{k\dot{\mu}\lambda}}. \tag{9}$$

Die Tatsache, daß die linearen Transformationen des Spinraumes für die a^k stets Lorentztransformationen, d. h. lineare Transformationen mit einer invarianten quadratischen Form der Signatur $- - - +$ induzieren, liegt daran, daß die Hermiteschen Matrices $a_{\dot{\lambda}\mu}$ eine Invariante derselben Signatur besitzen, nämlich ihre Determinante[1]

$$a_{11}\,a_{22} - a_{12}\,a_{21}$$

an deren Stelle wir auch den bis auf einen reellen Zahlenfaktor damit übereinstimmenden Ausdruck

$$a_{\dot{\lambda}\mu}\,a^{\dot{\lambda}\mu} = \gamma^{\dot{\lambda}\dot{\rho}}\,\gamma^{\mu\sigma}\,a_{\dot{\lambda}\mu}\,a_{\dot{\rho}\sigma} \tag{10}$$

betrachten können. Die Zuordnung $a_{\dot{\lambda}\mu} \rightarrow a^k$ verwandelt diese quadratische Form in eine quadratische Form in den a^k, von der wir verlangen, daß sie mit der metrischen Fundamentalform $g_{kl}\,a^k\,a^l$ übereinstimmt. Wir verlangen also

$$g_{kl}\,a^k\,a^l = g_{kl}\,\sigma^{k\dot{\lambda}\mu}\,\sigma^{l\dot{\rho}\sigma}\,a_{\dot{\lambda}\mu}\,a_{\dot{\rho}\sigma} = \gamma^{\dot{\lambda}\dot{\rho}}\,\gamma^{\mu\sigma}\,a_{\dot{\lambda}\mu}\,a_{\dot{\rho}\sigma}$$

identisch in $a_{\dot{\lambda}\mu}$; das heißt aber

$$g_{kl}\,\sigma^{k\dot{\lambda}\mu}\,\sigma^{l\dot{\rho}\sigma} = \gamma^{\dot{\lambda}\dot{\rho}}\,\gamma^{\mu\sigma}. \tag{11}$$

Beachtet man die Regeln für das Hinauf- und Herunterziehen der Indices, so kann man für (11) auch schreiben:

$$\sigma^{k\dot{\lambda}\mu}\,\sigma_{k\dot{\sigma}\rho} = \delta^{\dot{\lambda}}_{\dot{\sigma}}\,\delta^{\mu}_{\rho}$$

$$\left(\delta^{\dot{\lambda}}_{\dot{\sigma}} = \delta^{\mu}_{\sigma} = \begin{cases} +1 & \text{für } \dot{\lambda} = \dot{\sigma} \text{ und } \mu = \sigma \\ 0 & \text{»} \quad \dot{\lambda} \neq \dot{\sigma} \quad \text{»} \quad \mu \neq \sigma \end{cases}\right). \tag{12}$$

Auf Grund von (12) kann man (8) nach $a^{\dot{\lambda}\mu}$ auflösen:

$$a_{\dot{\lambda}\mu} = \sigma_{k\dot{\lambda}\mu}\,a^k.$$

---

[1] Um die Signatur dieser Form zu bestimmen, setze man $a_{11} = a + b$, $a_{22} = a - b$, $a_{12} = c + id$, $a_{21} = c - id$; dann wird $\Delta = a^2 - b^2 - c^2 - d^2$.

Setzt man diese Auflösung in (8) ein und vergleicht die Koeffizienten von a^k auf beiden Seiten, so folgt

$$\sigma^{k\dot{\lambda}\mu}\sigma_{l\dot{\lambda}\mu} = \delta^k_l . \tag{13}$$

Aus (13) folgt, wenn man auf beiden Seiten den Index l hochzieht:

$$g^{kl} = \sigma^{k\dot{\lambda}\mu}\sigma^l_{\dot{\lambda}\mu} . \tag{14}$$

Auf Grund der Formeln (12), (13) kann man nicht nur den Weltvektoren, sondern überhaupt allen Welttensoren umkehrbar eindeutig Spingrößen zuordnen, z. B. für einen Tensor p^{kl}:

$$\pi_{\dot{\lambda}\mu,\dot{\rho}\sigma} = \sigma_{k\dot{\lambda}\mu}\sigma_{l\dot{\rho}\sigma}p^{kl}; \quad p^{kl} = \sigma^{k\dot{\lambda}\mu}\sigma^{l\dot{\rho}\sigma}\pi_{\dot{\lambda}\mu,\dot{\rho}\sigma} .$$

Die Formel (14) zeigt noch einmal explizit, daß die Größen $\sigma^{k\dot{\lambda}\mu}$ und $\gamma_{\lambda\mu}$ zusammen das metrische Feld bestimmen. Das Umgekehrte gilt natürlich nicht.

Bei der oben besprochenen Zuordnung $\sigma_{\dot{\lambda}\mu} \to a^k$ entsprechen den zeitartigen Vektoren ($g_{kl}a^k a^l > 0$) solche Matrices $a_{\dot{\lambda}\mu}$, welche zu definiten Hermiteschen Formen gehören. Wir wollen die $\sigma^{k\dot{\lambda}\mu}$ immer so wählen, daß den in die Zukunft gerichteten Vektoren ($a^4 > 0$) die positiv-definiten Formen entsprechen. Weiter können wir die $\sigma^{k\dot{\lambda}\mu}$ stets so normieren, daß die rein-imaginäre 4reihige Determinante der Vektoren

$$\sigma^{k\dot{1}1}, \quad \sigma^{k\dot{1}2}, \quad \sigma^{k\dot{2}1}, \quad \sigma^{k\dot{2}2}$$

positiv-imaginär ausfällt. Sind diese Bedingungen nicht erfüllt, so multiplizieren wir die $\sigma^{k\dot{\lambda}\mu}$ mit -1 oder gehen zu den konjugiert-komplexen Werten über. Durch diese Normierungen wird erreicht, daß die $\sigma^{k\dot{\lambda}\mu}$ bis auf eine willkürliche Spintransformation durch das metrische Feld eindeutig bestimmt sind (vgl. § 5).

Der Ausdruck $\gamma_{\dot{\lambda}\dot{\rho}}(\sigma^{k\dot{\lambda}\mu}\sigma^{l\dot{\rho}\nu} + \sigma^{l\dot{\lambda}\mu}\sigma^{k\dot{\rho}\nu})$ ist offenbar antisymmetrisch in den Indices μ und ν, also gleich einem Vielfachen von $\gamma^{\mu\nu}$, etwa $h^{kl}\gamma^{\mu\nu}$. Multipliziert man noch mit $\gamma_{\nu\sigma}$ und beachtet die Regeln für das Herunterziehen von Indices, so folgt

$$\sigma^{k\dot{\lambda}\mu}\sigma^l_{\dot{\lambda}\sigma} + \sigma^{l\dot{\lambda}\mu}\sigma^k_{\dot{\lambda}\sigma} = h^{lk}\delta^\mu_\sigma .$$

Um h^{lk} zu bestimmen, setzen wir $\mu = \sigma$, summieren über μ und vergleichen mit (14). Man findet $h^{lk} = g^{lk}$. Damit ist die erste der beiden folgenden Formeln bewiesen:

$$\left.\begin{aligned} \sigma^{l\dot{\lambda}\mu}\sigma^k_{\dot{\lambda}\sigma} + \sigma^{k\dot{\lambda}\mu}\sigma^l_{\dot{\lambda}\sigma} &= g^{kl}\delta^\mu_\sigma \\ \sigma^{l\dot{\rho}\mu}\sigma^k_{\dot{\lambda}\mu} + \sigma^{k\dot{\rho}\mu}\sigma^l_{\dot{\lambda}\mu} &= g^{kl}\delta^{\dot{\rho}}_{\dot{\lambda}} \end{aligned}\right\} . \tag{15}$$

Die zweite folgt ganz analog.

Im ε-Formalismus muß man in (10) und (11) rechts $\gamma^{\dot\lambda\dot\rho}\gamma^{\mu\sigma}$ durch $\varepsilon^{\dot\lambda\dot\rho}\varepsilon^{\mu\sigma}$ ersetzen. Damit das Produkt $\varepsilon^{\dot\lambda\dot\rho}\varepsilon^{\mu\sigma}$ in (11) das richtige Gewicht erhält, muß man für $\sigma^{k\dot\lambda\rho}$ eine gemischte Dichte vom Absolutgewicht + 1 wählen. Es wird dann natürlich $\sigma_{k\dot\lambda\rho}$ eine gemischte Dichte vom Absolutgewicht − 1. Im ε-Formalismus wird nur der Spinordichte $a_{\dot\lambda\mu}$ vom Absolutgewichte − 1, oder der Spinordichte $a^{\dot\lambda\mu}$ vom Absolutgewichte + 1 in eineindeutiger Weise ein wirklicher Weltvektor zugeordnet, da nur dann $\sigma^{k\dot\lambda\mu}a_{\dot\lambda\mu}$ und $\sigma^{k}{}_{\dot\lambda\mu}\,a^{\dot\lambda\mu}$ bei einer Transformation im Spinraume keine Änderung erleidet.

Die Herleitung der Formeln (12) bis (15) bleibt im ε-Formalismus unverändert gültig.

§ 4. Der affine Zusammenhang.

Um invariante Differentialgleichungen für Spinoren aufstellen zu können, ist die Einführung einer kovarianten Differentiation (oder einer Parallelverschiebung) für Spinoren notwendig. Diese wird wie üblich für Spinvektoren durch lineare Formeln von der Art[1]

$$\left.\begin{aligned} \psi_{\alpha|k} &= \partial_k\psi_\alpha - \Gamma_{\alpha k}^{\rho}\psi_\rho \\ \psi^{\alpha}{}_{|k} &= \partial_k\psi^\alpha + \Gamma_{\rho k}^{\alpha}\psi^\rho \end{aligned}\right\} \qquad (16)$$

definiert. Transformiert man die Formeln (16) auf ein anderes Koordinatensystem und verlangt, daß die $\psi_{\alpha|k}$ oder $\psi^{\alpha}{}_{|k}$ sich kovariant, d. h. wie gemischte Größen transformieren sollen, so ergeben sich für Γ die Transformationsregeln

$$\Lambda_\sigma^\rho\,\Gamma_{\rho k}^{\alpha} = \Lambda_\rho^\alpha\,\Gamma'^{\rho}_{\sigma k} - \partial_k\Lambda_\sigma^\alpha .$$

Die Parallelverschiebung der konjugiert-komplexen Vektoren $\psi_{\dot\alpha}$, $\psi^{\dot\alpha}$ ist durch die der Vektoren ψ_α, ψ^α mitgegeben, sobald man verlangt, daß die Beziehung zwischen einem Vektor ψ^α und seinem konjugiert-komplexen Vektor $\psi^{\dot\alpha}$ bei Parallelverschiebung erhalten bleiben soll. Wir erhalten so eine kovariante Differentiation

$$\left.\begin{aligned} \psi_{\dot\alpha|k} &= \partial_k\psi_{\dot\alpha} - \Gamma_{\dot\alpha k}^{\dot\rho}\psi_{\dot\rho} \\ \psi^{\dot\alpha}{}_{|k} &= \partial_k\psi^{\dot\alpha} + \Gamma_{\dot\rho k}^{\dot\alpha}\psi^{\dot\rho} \end{aligned}\right\}$$

mit

$$\Gamma_{\dot\alpha k}^{\dot\beta} = \overline{\Gamma_{\alpha k}^{\beta}} .$$

Wir setzen weiter fest, daß ein beliebiger Tensor kovariant so differenziert werden soll wie ein Produkt von Vektoren, also z. B.:

$$a_{\dot\lambda\mu|k} = \partial_k a_{\dot\lambda\mu} - \Gamma_{\dot\lambda k}^{\dot\rho}a_{\dot\rho\mu} - \Gamma_{\mu k}^{\sigma}a_{\dot\lambda\sigma} .$$

---

[1] Zur Abkürzung ist $\partial_k = \dfrac{\partial}{\partial x^k}$ gesetzt. Wegen der Willkür der Koordinatenwahl im Spinraum hat die Differentiation ∂_k nicht kovarianten Charakter. Daraus ergibt sich die Notwendigkeit der Einführung der $\Gamma_{\alpha k}^{\rho}$. Daß man in den Formeln (16) beide Male dieselben $\Gamma_{\alpha k}^{\rho}$ mit entgegengesetzten Vorzeichen wählt, ist damit zu rechtfertigen, daß die invariante Relation $\psi_\alpha\chi^\alpha$ = konst. bei Parallelverschiebung der Vektoren ψ_α, χ^α erhalten bleiben soll, oder (was auf dasselbe hinauskommt) daß für ein Produkt $\psi_\alpha\chi^\alpha$ die Differentiationsregel gelten soll:

$$(\psi_\alpha\chi^\alpha)_{|s} = \psi_{\alpha|s}\chi^\alpha + \psi_\alpha\chi^{\alpha}{}_{|s} .$$

Wir verlangen von der kovarianten Differentiation, daß sie **volumtreu** sein soll, d. h. daß der Tensor $\gamma_{\alpha\beta}\gamma_{\dot{\alpha}\dot{\beta}}$ die kovariante Ableitung Null hat:

$$\gamma_{\alpha\beta|k}\gamma_{\dot{\alpha}\dot{\beta}} + \gamma_{\alpha\beta}\gamma_{\dot{\alpha}\dot{\beta}|k} = 0. \tag{17}$$

Die Bedingung (17) hat nur für die eine von Null verschiedene Komponente $\gamma_{12}\gamma_{\dot{1}\dot{2}}$ wirkliche Bedeutung; sie heißt dann[1]

$$\partial_k(\gamma_{12}\gamma_{\dot{1}\dot{2}}) - \Gamma^{\,1}_{1k}\gamma_{12}\gamma_{\dot{1}\dot{2}} - \Gamma^{\,2}_{2k}\gamma_{12}\gamma_{\dot{1}\dot{2}} - \Gamma^{\,\dot{1}}_{\dot{1}k}\gamma_{12}\gamma_{\dot{1}\dot{2}} - \Gamma^{\,\dot{2}}_{\dot{2}k}\gamma_{12}\gamma_{\dot{1}\dot{2}} = 0$$

oder, wenn $\gamma_{12}\gamma_{\dot{1}\dot{2}} = \gamma$ gesetzt wird,

$$\left.\begin{aligned} \partial_k\gamma - (\Gamma^{\,\alpha}_{\alpha k} + \Gamma^{\,\dot{\alpha}}_{\dot{\alpha}k})\gamma &= 0 \\ \Gamma^{\,\alpha}_{\alpha k} + \Gamma^{\,\dot{\alpha}}_{\dot{\alpha}k} &= \partial_k \log\gamma \end{aligned}\right\}. \tag{18}$$

Da auf Grund von (8) mit jedem Tensor $a_{\lambda\mu}$ ein Weltvektor a^k verbunden ist, so induziert die Parallelverschiebung der Tensoren $a_{\lambda\mu}$ zugleich eine Parallelverschiebung der Vektoren a^k. Die zugehörige kovariante Ableitung ist durch

$$a^k{}_{|l} = \sigma^{k\lambda\mu}a_{\lambda\mu|l} \tag{19}$$

definiert. Man erhält dasselbe Ergebnis, wenn man die kovariante Differentiation von a^k in der üblichen Weise durch ein $\Gamma^{\,s}_{kl}$ festlegt und dieses $\Gamma^{\,s}_{kl}$ so bestimmt, daß die kovariante Ableitung von $\sigma^{k\lambda\mu}$ gleich Null wird:

$$\sigma^{k\lambda\mu}{}_{|s} = \partial_s\sigma^{k\lambda\mu} + \Gamma^{\,k}_{rs}\sigma^{r\lambda\mu} + \Gamma^{\,\lambda}_{\rho s}\sigma^{k\rho\mu} + \Gamma^{\,\mu}_{\sigma s}\sigma^{k\lambda\sigma} = 0. \tag{20}$$

Die Formel (20) bestimmt offenbar die $\Gamma^{\,k}_{rs}$ eindeutig und hat (19) zur Folge. Eine weitere Folge von (17), (20) und der Differentiationsregel für Produkte ist, daß die kovariante Ableitung von

$$g^{kl} = \sigma^{k\lambda\mu}\sigma^{l}{}_{\lambda\mu} = \sigma^{k\lambda\mu}\sigma^{l\rho\sigma}\gamma_{\lambda\rho}\gamma_{\mu\sigma}$$

gleich Null wird[2]:

$$g^{kl}{}_{|s} = 0. \tag{21}$$

Wenn man nun noch verlangt, daß die Übertragung $\Gamma^{\,s}_{kl}$ symmetrisch ausfällt:

$$\Gamma^{\,k}_{rs} = \Gamma^{\,k}_{sr}, \tag{22}$$

so folgt aus (21) bekanntlich, daß die $\Gamma^{\,k}_{rs}$ gleich den durch g_{kl} bestimmten Christoffelsymbolen $\left\{ {rs \atop k} \right\}$ sein müssen. Die Symmetrieannahme ist die einfachste, die man machen kann: würde man sie nicht machen, so würde man als Differenz $\Gamma^{\,k}_{rs} - \Gamma^{\,k}_{sr}$ einen Tensor $S_{rs}^{\,k}$ erhalten, dessen physikalische Bedeutung nicht bekannt ist.

---

[1] Die Gleichung (17) ließe sich auch (wie später gezeigt wird) aus der Forderung $g_{kl|s} = 0$ ableiten, d. h. aus der Bedingung, daß ein starrer Maßstab bei Parallelverschiebung seine Länge nicht ändert.

[2] Man hätte auch umgekehrt von der Forderung (21) ausgehen und daraus umgekehrt (17) und (18) herleiten können.

Die Symmetriebedingung (22) ist, wenn man die Γ^k_{rs} aus (20) berechnet, mit 6.4 linearen Bedingungen für die 8.4 Real- und Imaginärteile der 4.4 Komponenten $\Gamma^{\mu}_{\rho s}$ gleichwertig. Diese reichen, zusammen mit den 4 Bedingungen (18), noch nicht zur Bestimmung der $\Gamma^{\mu}_{\rho s}$ aus, sondern es bleiben noch 4 reelle Parameter willkürlich. In der Tat: Sind die Gleichungen (18), (20) erfüllt, so bleiben sie es, wenn man die Γ^k_{rs} ungeändert läßt, die $\Gamma^{\mu}_{\rho s}$ und $\Gamma^{\dot\mu}_{\dot\rho s}$ aber durch

$$\Gamma^{\mu}_{\rho s} + \frac{i}{2}\varphi_s \delta^{\mu}_{\rho} \quad \text{und} \quad \Gamma^{\dot\mu}_{\dot\rho s} - \frac{i}{2}\varphi_s \delta^{\dot\mu}_{\dot\rho}$$

ersetzt. Es tritt hier also ganz von selbst eine willkürliche reelle Größe Φ_s auf, die durch

$$\Gamma^{\alpha}_{\alpha s} - \Gamma^{\dot\alpha}_{\dot\alpha s} = 2i\Phi_s \tag{23}$$

definiert wird. Wir wollen zunächst die Transformationseigenschaften von Φ_s untersuchen. Gegenüber Raumtransformationen verhält sich Φ_s wie ein Weltvektor. Bei Transformationen im Spinraume, wenn $\Delta = |\Delta| e^{i\varphi}$ ist, erhält man dagegen:

$$\Phi'_s = \Phi_s - \partial_s \varphi\,.$$

Dies folgt aus den Transformationsgleichungen

$$\Gamma'^{\alpha}_{\alpha s} = \Gamma^{\alpha}_{\alpha s} - \partial_s \log \Delta \quad \text{und} \quad \Gamma'^{\dot\alpha}_{\dot\alpha s} = \Gamma^{\dot\alpha}_{\dot\alpha s} - \partial_s \log \bar\Delta\,.$$

Die Φ_s transformieren sich daher bei einer Transformation von der Form $a'^{\lambda} = e^{\frac{1}{2}i\varphi} a^{\lambda}$ genau so, wie es Potentiale nach dem Weylschen Prinzip der Eichinvarianz tun sollen. Φ_s soll daher mit dem elektromagnetischen Potential identifiziert werden.

Aus der Gleichung $\gamma_{12} = \sqrt{\gamma}\, e^{i\vartheta}$ folgt, daß sich $\partial_s\vartheta$ genau so wie Φ_s transformiert, d. h.

$$\partial_s\vartheta' = \partial_s\vartheta - \partial_s\varphi\,.$$

Daher ist

$$\Phi^*_s = \Phi_s - \partial_s\vartheta$$

ein wirklicher Vektor. Er ist vom elektromagnetischen Potential nur um den Gradienten einer Funktion verschieden. Eine physikalische Bedeutung scheint der Vektor Φ^*_s nicht zu haben.

Aus (18) und (23) lassen sich die verjüngten Größen $\Gamma^{\alpha}_{\alpha s}$ und $\Gamma^{\dot\alpha}_{\dot\alpha s}$ explizite berechnen:

$$\left.\begin{aligned} \Gamma^{\alpha}_{\alpha s} &= i\Phi_s + \partial_s \log\sqrt{\gamma} = i\Phi^*_s + \partial_s \log\sqrt{\gamma}\, e^{i\vartheta} \\ \Gamma^{\dot\alpha}_{\dot\alpha s} &= -i\Phi_s + \partial_s \log\sqrt{\gamma} = -i\Phi^*_s + \partial_s \log\sqrt{\gamma}\, e^{-i\vartheta} \end{aligned}\right\}.$$

Für kovariante Ableitung des Spinors $\gamma_{\lambda\mu}$ findet man daraus:

$$\left.\begin{aligned} \gamma^{\lambda\mu}_{|s} &= i\gamma^{\lambda\mu}\,\Phi^*_s & \gamma^{\dot\lambda\dot\mu}_{|s} &= -i\gamma^{\dot\lambda\dot\mu}\,\Phi^*_s \\ \gamma_{\lambda\mu|s} &= -i\gamma_{\lambda\mu}\,\Phi^*_s & \gamma_{\dot\lambda\dot\rho|s} &= i\gamma_{\dot\lambda\dot\mu}\,\Phi^*_s \end{aligned}\right\}. \tag{24}$$

Im ε-Formalismus kann man die obigen Formeln (17) bis (24) nicht ohne weiteres übernehmen, sondern man muß zuerst eine kovariante Ableitung für Spinordichten definieren. Das geschieht dadurch, daß man zuerst für die kovariante Differentiation der Spindichten vom Gewicht $\mathfrak{f}$ den Ansatz

$$a_k = \partial_k a - \mathfrak{f}\,\Gamma_k\, a \tag{25}$$

macht. Der Faktor $\mathfrak{f}$ wird darum hinzugefügt, weil eine Dichte vom Gewicht $\mathfrak{f}$ die $\mathfrak{f}$-te Potenz einer Dichte vom Gewicht Eins ist. Damit das so definierte a_k eine Vektordichte sei, müssen die Γ_k sich bei einer Raumtransformation wie Vektorkomponenten, bei einer Spintransformation dagegen nach der Formel

$$\Gamma'_k = \Gamma_k - \partial_k \log \Delta \tag{26}$$

transformieren. Die kovariante Differentiation von beliebigen Spinordichten wird nun so definiert, wie sie sich für Produkte von Spinoren mit skalaren Dichten von selbst ergibt, also nach der gewöhnlichen Formel für kovariante Spinordifferentiation mit einem Zusatzglied $-\mathfrak{f}\,\Gamma_k$ mal dem betreffenden Spinor.

Die zunächst beliebig eingeführten Γ_k lassen sich eindeutig festlegen durch die invariante Forderung, daß die kovariante Ableitung der Spinordichte $\varepsilon_{\lambda\mu}$ (vom Gewicht -1) gleich Null ausfallen soll:

$$\left.\begin{aligned} \varepsilon_{12|k} = -\Gamma^1_{1k}\varepsilon_{12} - \Gamma^2_{2k}\,\varepsilon_{12} + \Gamma_k\,\varepsilon_{12} = 0 \\ \Gamma_k = \Gamma^a_{ak} \end{aligned}\right\}. \tag{27}$$

Für die konjugiert-komplexen Dichten vom Gewicht $\dot{\mathfrak{f}}$ ist in den obigen Formeln überall Γ_k durch das konjugiert-komplexe $\overline{\Gamma}_k$ zu ersetzen.

Das Produkt zweier Dichten der Gewichte $\mathfrak{f}$, $\dot{\mathfrak{f}}$ ist eine Dichte vom Absolutgewicht $2\mathfrak{f}$, deren kovariante Ableitung nach der gewöhnlichen Produkt-Differentiationsregel so heißen würde:

$$a_k = \partial_k a - \mathfrak{f}\,(\Gamma_k + \overline{\Gamma}_k)\,a\,.$$

Wir setzen nun

$$\Gamma_k + \overline{\Gamma}_k = \Gamma^a_{ak} + \Gamma^{\dot a}_{\dot a k} = 2\,\Pi_k$$

und definieren für beliebige Dichten vom Absolutgewicht $\mathfrak{f}$ die kovariante Ableitung durch

$$a_k = \partial_k a - \mathfrak{f}\,\Pi_k a\,.$$

Wie man Spinordichten vom Absolutgewicht $\mathfrak{f}$ und gemischte Dichten zu differenzieren hat, ist nunmehr klar.

Aus denselben Gründen wie oben werden wir versuchen, die kovariante Differentiation der Spinoren und der Weltvektoren so aufeinander abzustimmen, daß die kovariante Ableitung von $\sigma^{k\dot\lambda\mu}$ gleich Null ausfällt. Das gibt, weil $\sigma^{k\dot\lambda\mu}$ im ε-Formalismus eine Größe vom Absolutgewicht 1 ist, die Bedingung

$$\sigma^{k\dot\lambda\mu}{}_{|s} = \partial_s\,\sigma^{k\dot\lambda\mu} + \Gamma^k_{rs}\,\sigma^{r\dot\lambda\mu} + \Gamma^{\dot\lambda}_{\dot\rho s}\,\sigma^{k\dot\rho\mu} + \Gamma^{\mu}_{\sigma s}\,\sigma^{k\dot\lambda\sigma} - \Pi_s\,\sigma^{k\dot\lambda\mu} = 0\,. \tag{28}$$

welche wieder die Γ^k_{rs} durch die $\Gamma^{\mu}_{\sigma s}$ und $\Gamma^{\dot\lambda}_{\dot\rho s}$ auszudrücken gestatten. Die Forderung, daß Γ^k_{rs} symmetrisch ausfällt, ergibt 6.4 Bedingungen für die 8.4 Real- und Imaginärteile der $\Gamma^{\mu}_{\rho s}$. Es bleiben also jetzt 8 reelle Parameter willkürlich. In der Tat: (28) bleibt erfüllt, wenn man $\Gamma^{\mu}_{\rho s}$ und $\Gamma^{\dot\mu}_{\dot\rho s}$ durch

$$\Gamma^{\mu}_{\rho s} + \tfrac{1}{2}(\pi_s + i\,\varphi_s)\,\delta^{\mu}_{\rho} \quad \text{und} \quad \Gamma^{\dot\mu}_{\dot\rho s} + \tfrac{1}{2}(\pi_s - i\,\varphi_s)\,\delta^{\dot\mu}_{\dot\rho}$$

ersetzt. Setzen wir also

$$\Gamma_s = \Gamma^{\alpha}_{\alpha s} = \Pi_s + i\Phi_s\,; \quad \overline{\Gamma}_s = \Gamma^{\dot\alpha}_{\dot\alpha s} = \Pi_s - i\Phi_s\,, \tag{29}$$

so finden wir, daß in diesem Formalismus nicht nur die Φ_s, sondern auch die Π_s willkürlich sind.

Wir werden jedoch sehen, daß diese letztere Willkür aus den physikalisch wichtigen Formeln ganz herausfällt, weil in diesen Formeln die $\Gamma^{\mu}_{\rho s}$ immer nur in den Kombinationen

$$\Gamma^{\mu}_{\rho s} - \tfrac{1}{2}\,\delta^{\mu}_{\rho}\,\Pi_s \quad \text{und} \quad \Gamma^{\dot\mu}_{\dot\rho s} - \tfrac{1}{2}\,\delta^{\dot\mu}_{\dot\rho}\,\Pi_s$$

vorkommen. Diese Kombinationen sind nämlich für die kovariante Differentiation der kontravarianten Spinvektordichten vom Gewicht $\frac{1}{2}$ und der kovarianten Spinvektordichten vom Gewicht $-\frac{1}{2}$ maßgebend, und es zeigt sich (vgl. § 7), daß nur solche Vektordichten eine wirkliche Rolle spielen. Wesentlich tritt aber in den Formeln die Größe Φ_s auf, die sich auch in diesem Formalismus bei Raumtransformationen wie ein Vektor, bei Spintransformationen von der Determinante $\Delta = |\Delta|\, e^{i\varphi}$ aber nach der aus (26) und (29) folgenden Formel

$$\Phi'_s = \Phi_s - \partial_s\varphi \tag{30}$$

transformiert.

Diese Größe Φ_s soll wieder mit dem elektromagnetischen Potential identifiziert werden.

§ 5. Die vollständig geodätischen Systeme.

Wir wollen jetzt die Komponenten der Größe $\sigma^{k\lambda\dot\mu}$ und der Parallelverschiebung in der Umgebung eines Weltpunktes P_0 in einem passend gewählten Koordinatensystem berechnen.

Ist der Raum euklidisch, so genügt man in einem in der Welt und im Spinraum passend gewählten Koordinatensystem den Bedingungen (11), (18), (20) und (22) durch folgende Werte der auftretenden Größen:

1. $g_{kl} = \mathring{g}_{kl}\,; \quad \mathring{g}_{11} = \mathring{g}_{22} = \mathring{g}_{33} = -\mathring{g}_{44} = -1\,,$ sonst $\mathring{g}_{kl} = 0$ (31a)

2. $\gamma^{\lambda\mu} = \varepsilon^{\lambda\mu}\,; \quad \varepsilon^{12} = -\varepsilon^{21} = 1\,, \quad \varepsilon^{11} = \varepsilon^{22} = 0$ (31b)

3. $\sigma^{k\lambda\dot\mu} = \mathring{\sigma}^{k\lambda\dot\mu}\,; \quad \mathring{\sigma}^{1\lambda\dot\mu} = \dfrac{1}{\sqrt{2}}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad \mathring{\sigma}^{2\lambda\dot\mu} = \dfrac{1}{\sqrt{2}}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$

$\qquad \mathring{\sigma}^{3\lambda\dot\mu} = \dfrac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}; \quad \mathring{\sigma}^{4\lambda\dot\mu} = \dfrac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ (31c)

4. $\Gamma^{l}_{sk} = \left\{ {sk \atop l} \right\} = 0$ (31d)

5. $\Gamma^{\beta}_{\alpha s} = \tfrac{1}{2}\, i\,\Phi_s\,\delta^{\beta}_{\alpha}\,.$ (31e)

Diese Werte bleiben erhalten, wenn man in der Welt eine beliebige Lorentz-Transformation und gleichzeitig im Spinraum eine passende konstante lineare Transformation von der Determinante Eins ausführt[1].

[1] Für den Beweis siehe etwa B. L. v. d. Waerden, Die Gruppentheoretische Methode in der Quantenmechanik, § 20.

Bei gegebenen $g_{kl} = \mathring{g}_{kl}$ und $\gamma^{\lambda\mu} = \varepsilon^{\lambda\mu}$ sind die in (31c) angegebenen $\sigma^{k\dot{\lambda}\mu}$ nicht die einzigen, welche die Bedingungen (11) erfüllen, sondern man hat noch die Freiheit einer willkürlichen Lorentz-Transformation, auszuüben auf den Index k, oder auch einer willkürlichen Spintransformation von der Determinante Eins, auszuüben auf die Indices $\dot{\lambda}$, μ. Das ist aber auch die einzige Freiheit, die man hat. Denn die Bedingung (11) besagt laut Herleitung, daß bei der Zuordnung $a_{\dot{\lambda}\mu} \to a^k$ die quadratische Form $\gamma^{\dot{\lambda}\dot{\rho}}\gamma^{\mu\sigma}a_{\dot{\lambda}\mu}a_{\dot{\rho}\sigma}$ in $\mathring{g}_{kl}a^k a^l$ übergeführt wird, und eine solche lineare Zuordnung ist völlig bestimmt bis auf eine lineare Transformation, welche die letztere quadratische Form invariant läßt, d. h. aber bis auf eine Lorentz-Transformation[1].

Wir gehen jetzt zum allgemeinen Fall des Riemannschen Raumes über. Wir denken uns, falls die γ-Symbolik zugrunde gelegt wird, das Koordinatensystem im Spinraum immer so gewählt, daß $\gamma^{12} = 1$ wird. Dann ist also zwischen $\gamma^{\lambda\mu}$ und $\varepsilon^{\lambda\mu}$ kein Unterschied. Für einen einzelnen Weltpunkt P_0 kann man außerdem durch eine passende lineare Transformation

$$x^k = c^k_{\ p}\,\mathring{x}^p,$$

die g_{kl} in $\mathring{g}_{kl}$ überführen. Die Vektoren $c^k_{\ 1}$, $c^k_{\ 2}$, $c^k_{\ 3}$, $c^k_{\ 4}$ müssen dazu nur ein orthogonales n-Bein bilden. Dann kann man durch den Ansatz

$$\sigma^{k\dot{\lambda}\mu} = c^k_{\ p}\,\mathring{\sigma}^{p\dot{\lambda}\mu}$$

ein System von Komponenten $\sigma^{k\dot{\lambda}\mu}$ bestimmen, welches die Gleichung (11) befriedigt, also zum metrischen Feld paßt. Dieses System $\sigma^{k\dot{\lambda}\mu}$ ist nach dem oben Bemerkten auch im wesentlichen (d. h. bis auf eine lineare Spintransformation) das einzige. Läßt man die $c^k_{\ p}$ differenzierbar vom betrachteten Weltpunkt abhängen, so ist das entstehende $\sigma^{k\dot{\lambda}\mu}$-Feld auch differenzierbar.

Man kann aber für die unmittelbare Umgebung eines Punktes P_0 noch mehr erreichen. Wählt man das Koordinatensystem in P_0 so, daß in diesem Punkt $g_{kl} = \mathring{g}_{kl}$ wird, so kann man in diesem Punkt auch $c^k_p = \delta^k_p$, $\sigma^{k\dot{\lambda}\mu} = \mathring{\sigma}^{k\dot{\lambda}\mu}$ wählen. Nun kann man für jeden Punkt P der unmittelbaren Umgebung von P_0 das orthogonale n-Bein $c^k_{\ p}$ durch Parallelverschiebung des n-Beins δ^k_p von P_0 nach P bestimmen und die $\sigma^{k\dot{\lambda}\mu}$ entsprechend festlegen. Analytisch läßt sich diese Parallelverschiebung dadurch leicht verfolgen, daß man ein in P_0 geodätisches Koordinatensystem zugrunde legt, in bezug auf welches also

$$(\partial_s g_{kl})_{P_0} = 0, \quad (\Gamma^{\ l}_{sk})_{P_0} = 0$$

ist. Die parallel verschobenen Vektoren $c^k_{\ p}$ genügen dann einfach der Bedingung

$$(\partial_s c^k_{\ p})_{P_0} = 0.$$

---

[1] Die »uneigentlichen« Lorentz-Transformationen, welche die Ablaufsrichtung der Zeit oder die Orientierung des Raumes umkehren, können auf Grund der in § 3 getroffenen Normierungen hier nicht auftreten.

Daraus folgt dann sofort

$$(\partial_s \sigma^{k\dot\lambda\mu})_{P_0} = 0,$$

d. h.: die $\sigma^{k\dot\lambda\mu}$ sind wie die g_{kl} und $\gamma^{\lambda\mu}$ in bezug auf das gewählte Koordinatensystem in erster Näherung konstant. Aus dem so gefundenen $\sigma^{k\dot\lambda\mu}$-Feld entsteht das allgemeinste bei den einmal gewählten Raumkoordinaten durch eine lineare Spintransformation.

Wir wollen ein solches Koordinatensystem, dessen Existenz eben bewiesen wurde, in bezug auf welches die Bedingungen

$$(g_{kl})_{P_0} = \overset{\circ}{g}_{kl}, \qquad (\partial_s g_{kl})_{P_0} = 0 \tag{32a}$$

$$(\gamma_{\lambda\mu})_{P_0} = \varepsilon_{\lambda\mu}, \qquad (\partial_s \gamma_{\lambda\mu})_{P_0} = 0 \tag{32b}$$

$$(\sigma^{k\dot\lambda\mu})_{P_0} = \overset{\circ}{\sigma}{}^{k\dot\lambda\mu}, \quad (\partial_s \sigma^{k\dot\lambda\mu})_{P_0} = 0 \tag{32c}$$

erfüllt sind, ein vollständig geodätisches Koordinatensystem in P_0 nennen.

Wir können nun in bezug auf ein vollständig geodätisches Koordinatensystem in P_0 die Komponenten $\Gamma^{\rho}_{\alpha k}$ der Parallelverschiebung der Spinoren berechnen. Die Gleichungen (20) und (23) nehmen in P_0 die folgende einfache Form an:

$$\Gamma^{\dot\lambda}_{\dot\rho s} \overset{\circ}{\sigma}{}^{k\dot\rho\mu} + \Gamma^{\mu}_{\rho s} \overset{\circ}{\sigma}{}^{k\dot\lambda\rho} = 0 \tag{33}$$

$$\Gamma^{\dot\alpha}_{\dot\alpha s} - \Gamma^{\alpha}_{\alpha s} = -2i\,\Phi_s. \tag{34}$$

Dieses elementare lineare Gleichungssystem hat nur eine Lösung; nämlich:

$$(\Gamma^{\beta}_{\alpha s})_{P_0} = \tfrac{1}{2} i\, \delta^{\beta}_{\alpha}\, \Phi_s. \tag{35}$$

Damit ist der bisher noch ausstehende Beweis geliefert, daß die $\Gamma^{\beta}_{\alpha s}$ bis auf die Willkür der Φ_s durch die Gleichungen (20) wirklich eindeutig bestimmt sind. Wenn das nämlich im vollständig geodätischen Koordinatensystem gilt, gilt es auch in jedem anderen.

Im ε-Formalismus fällt die Bedingung (32b) fort, und das hat zur Folge, daß ein vollständig geodätisches Bezugssystem noch die willkürliche Spintransformation $\alpha'^{\lambda} = \beta\,\alpha^{\lambda}$, $\beta = \varrho e^{i\vartheta}$, zuläßt.

Statt (20) hat man jetzt die Gleichung (28), die aus (20) entsteht, indem man $\Gamma^{\mu}_{\sigma s}$ und $\Gamma^{\dot\lambda}_{\dot\rho s}$ durch $\Gamma^{\mu}_{\sigma s} - \frac{1}{2}\Pi_s\,\delta^{\mu}_{\sigma}$ und $\Gamma^{\dot\lambda}_{\dot\rho s} - \frac{1}{2}\Pi_s\,\delta^{\dot\lambda}_{\dot\rho}$ ersetzt. Macht man diese Ersetzung auch in der Lösung (35), so erhält man die allgemeine Lösung im vollständig geodätischen Koordinatensystem:

$$(\Gamma^{\mu}_{\sigma s})_{P_0} = \tfrac{1}{2}(\Pi_s + i\,\Phi_s)\,\delta^{\mu}_{\sigma}$$

$$(\Gamma^{\dot\lambda}_{\dot\rho s})_{P_0} = \tfrac{1}{2}(\Pi_s - i\,\Phi_s)\,\delta^{\dot\lambda}_{\dot\rho}.$$

Auch hier besteht also außer der Willkür der Π_s und der Φ_s keine weitere Willkür bei der Bestimmung der $\Gamma^{\mu}_{\rho s}$. Die Willkür der Π_s fällt, wie schon bemerkt, heraus, wenn statt der $\Gamma^{\mu}_{\rho s}$ und $\Gamma^{\dot\mu}_{\dot\rho s}$ nur die Größen

$$\overset{*}{\Gamma}{}^{\mu}_{\rho s} = \Gamma^{\mu}_{\rho s} - \tfrac{1}{2}\delta^{\mu}_{\rho}\,\Pi_s \quad \text{und} \quad \overset{*}{\Gamma}{}^{\dot\mu}_{\dot\rho s} = \Gamma^{\dot\mu}_{\dot\rho s} - \tfrac{1}{2}\delta^{\mu}_{\rho}\,\Pi_s, \tag{36}$$

die für die kovariante Differentiation der kontra-(ko-)varianten Spinvektordichten vom Gewicht $\frac{1}{2}\,(-\frac{1}{2})$ maßgebend sind, betrachtet werden.

§ 6. Die Krümmungstensoren.

Im Riemannschen Raum existiert der wohlbekannte Krümmungstensor:

$$R^{r}_{kps} = -\partial_s \Gamma^{r}_{kp} + \partial_p \Gamma^{r}_{ks} - \Gamma^{h}_{kp} \Gamma^{r}_{hs} + \Gamma^{h}_{ks} \Gamma^{r}_{hp}.$$

Ähnlich kann auch für den Spinraum der gemischte Krümmungstensor gebildet werden. Seine Komponenten sind:

$$\left.\begin{aligned} P^{\mu}_{\lambda ps} &= -\partial_s \Gamma^{\mu}_{\lambda p} + \partial_p \Gamma^{\mu}_{\lambda s} - \Gamma^{\rho}_{\lambda p} \Gamma^{\mu}_{\rho s} + \Gamma^{\mu}_{\rho p} \Gamma^{\rho}_{\lambda s} \\ P^{\dot\mu}_{\dot\lambda ps} &= -\partial_s \Gamma^{\dot\mu}_{\dot\lambda p} + \partial_p \Gamma^{\dot\mu}_{\dot\lambda s} - \Gamma^{\dot\rho}_{\dot\lambda p} \Gamma^{\dot\mu}_{\dot\rho s} + \Gamma^{\dot\mu}_{\dot\rho p} \Gamma^{\dot\rho}_{\dot\lambda s} \end{aligned}\right\}. \tag{37}$$

Durch Verjüngung erhält man mit Rücksicht auf (23a) den schiefsymmetrischen Welttensor:

$$\left.\begin{aligned} P^{\rho}_{\rho ps} &= i(\partial_p \Phi_s - \partial_s \Phi_p) = iF_{ps} \\ P^{\dot\rho}_{\dot\rho ps} &= -i(\partial_p \Phi_s - \partial_s \Phi_p) = -iF_{ps} \end{aligned}\right\}. \tag{38}$$

Es tritt hier von selbst der Welttensor $F_{ps} = -F_{sp}$ der **elektromagnetischen Feldstärke** auf.

Es soll jetzt die zwischen dem gemischten und Riemannschen Krümmungstensor bestehende Beziehung aufgesucht werden. Man findet zunächst durch Ausrechnung die gewohnten Formeln:

$$\left.\begin{aligned} \psi^{\rho}{}_{|kl} - \psi^{\rho}{}_{|lk} &= \psi^{\sigma} P^{\rho}_{\sigma lk} \\ \psi_{\dot\rho|kl} - \psi_{\dot\rho|lk} &= \psi_{\dot\sigma} P^{\dot\sigma}_{\dot\rho kl} \end{aligned}\right\} \tag{39a}$$

$$\sigma^{k\dot\lambda\mu}{}_{|ps} - \sigma^{k\dot\lambda\mu}{}_{|sp} = \sigma^{k\dot\rho\mu} P^{\dot\lambda}_{\dot\rho sp} + \sigma^{k\dot\lambda\rho} P^{\mu}_{\rho sp} + \sigma^{r\dot\lambda\mu} R^{k}_{rsp}. \tag{39b}$$

Die linke Seite der Gleichung (39b) verschwindet identisch infolge von (20). Man hat also:

$$\sigma^{k\dot\rho\mu} P^{\dot\lambda}_{\dot\rho sp} + \sigma^{k\dot\lambda\rho} P^{\mu}_{\rho sp} + \sigma^{r\dot\lambda\mu} R^{k}_{rsp} = 0. \tag{40}$$

Man kann ohne weiteres R^{k}_{rsp} durch $P^{\dot\lambda}_{\dot\rho sp}$ und $P^{\lambda}_{\rho sp}$ ausdrücken. Man kann aber auch umgekehrt R^{k}_{rsp} als gegeben annehmen und die $P^{\lambda}_{\rho sp}$ aus (40) und (38) zu bestimmen versuchen. Die Gleichungen (38) und (40) sind befriedigt, wenn man setzt

$$\left.\begin{aligned} P^{\lambda}_{\rho sp} &= \tfrac{1}{2} R_{krsp} \sigma^{k\lambda\dot\nu} \sigma^{r}_{\dot\nu\rho} + \tfrac{1}{2} iF_{sp} \delta^{\lambda}_{\rho} \\ P^{\dot\lambda}_{\dot\rho sp} &= \tfrac{1}{2} R_{krsp} \sigma^{k\dot\lambda\nu} \sigma^{r}_{\nu\dot\rho} - \tfrac{1}{2} iF_{sp} \delta^{\dot\lambda}_{\dot\rho} \end{aligned}\right\}. \tag{41}$$

Eine andere Lösung kann es bei gegebenen R^{k}_{rsp} und F_{sp} nicht geben, denn die Differenz zweier Lösungen müßte dem Gleichungssystem

$$\sigma^{k\dot\rho\mu} P^{\dot\lambda}_{\dot\rho sp} + \sigma^{k\dot\lambda\rho} P^{\mu}_{\rho sp} = 0$$

$$P^{\rho}_{\rho sp} = 0, \quad P^{\dot\rho}_{\dot\rho sp} = 0$$

genügen, welches für feste p (und $\sigma^{k\dot{\lambda}\mu} = \mathring{\sigma}^{k\dot{\lambda}\mu}$) mit dem Gleichungssystem (33), (34) (mit $\Phi_s = 0$) übereinstimmt, von welchem wir schon feststellten, daß es nur die Nullösung besitzt.

Im ε-Formalismus kann man sowohl aus den $\Gamma^{\mu}_{\lambda s}$ als auch aus den durch (36) definierten $\overset{*}{\Gamma}{}^{\mu}_{\lambda s}$ einen Krümmungstensor bilden. Tut man das letztere, ersetzt also in (37) überall $\Gamma^{\mu}_{\lambda s}$ durch $\overset{*}{\Gamma}{}^{\mu}_{\lambda s}$, so gelten auch die Formeln (38). Wir wollen uns für diese letztere Möglichkeit entscheiden, da die Parallelverschiebung $\Gamma^{\mu}_{\lambda s}$, in deren Krümmung noch die unbestimmte Größe Π_s eingeht, keine physikalische Bedeutung zu haben scheint. Die Formeln (39) gelten auch, sofern man unter ψ^{ρ} und $\psi_{\dot{\beta}}$ Spinordichten von den Absolutgewichten $\frac{1}{2}$ und $-\frac{1}{2}$ versteht. Ebenso gelten (40) und (41).

§ 7. Die Diracschen Gleichungen.

Den Ausgangspunkt unserer Betrachtungen bildet der Erhaltungssatz. Es sei $\mathfrak{J}^k$ der Stromvektor, der dem Erhaltungssatze gehorchen soll. Im Riemannschen Kontinuum ist daher:

$$\frac{1}{\sqrt{-g}}\,\partial_k \sqrt{-g}\,\mathfrak{J}^k = \mathfrak{J}^k{}_{|k} = 0.$$

Der Weltvektor $\mathfrak{J}^k$ induziert einen Spintensor $\varkappa^{\dot{\lambda}\mu}$, wobei

$$\mathfrak{J}^k = \sigma^{k\dot{\lambda}\mu}\varkappa_{\dot{\lambda}\mu} = \sigma^k{}_{\dot{\lambda}\mu}\varkappa^{\dot{\lambda}\mu}$$

ist. Wir verlangen also

$$\sigma^{k\dot{\lambda}\mu}\varkappa_{\dot{\lambda}\mu|k} = \sigma^k{}_{\dot{\lambda}\mu}\varkappa^{\dot{\lambda}\mu}{}_{|k} = 0. \tag{42}$$

Da $\mathfrak{J}^k$ ein reeller Weltvektor ist, muß $\varkappa^{\dot{\lambda}\mu}$ in den Indices $\dot{\lambda}, \mu$ vertauschbar sein. Dieser Bedingung und der anderen, daß $\mathfrak{J}^k$ ein zeitartiger Vektor mit $\mathfrak{J}^4 \geq 0$ sein soll, genügt man durch den folgenden einfachen Ansatz:

$$\varkappa^{\dot{\lambda}\mu} = \psi^{\dot{\lambda}}\psi^{\mu} + \chi^{\dot{\lambda}}\chi^{\mu}, \tag{43}$$

wo ψ^{λ}, χ^{μ} die Komponenten zweier Spinvektoren bezeichnen. Gleichung (42) kann in der folgenden Form geschrieben werden:

$$(\sigma^k{}_{\dot{\lambda}\mu}\psi^{\dot{\lambda}}\psi^{\mu} + \sigma^{k\dot{\lambda}\mu}\chi_{\dot{\lambda}}\chi_{\mu})_{|k} = 0. \tag{44}$$

Durch Entwicklung dieser Gleichung findet man

$$\left.\begin{aligned} \psi^{\mu}\sigma^k{}_{\dot{\lambda}\mu}\psi^{\dot{\lambda}}{}_{|k} + \chi_{\mu}\sigma^{k\dot{\lambda}\mu}\chi_{\dot{\lambda}|k} & \\ + \psi^{\dot{\lambda}}\sigma^k{}_{\dot{\lambda}\mu}\psi^{\mu}{}_{|k} + \chi_{\dot{\lambda}}\sigma^{k\dot{\lambda}\mu}\chi_{\mu|k} &= 0 \end{aligned}\right\}. \tag{45}$$

Der Erhaltungssatz ist daher erfüllt, wenn:

$$\left.\begin{aligned} \sigma^k{}_{\dot{\lambda}\mu}\psi^{\dot{\lambda}}{}_{|k} &= \alpha\chi_{\mu} \\ \sigma^{k\dot{\lambda}\mu}\chi_{\dot{\lambda}|k} &= -\alpha\psi^{\mu} \end{aligned}\right\}. \tag{46a}$$

Die zu diesen Gleichungen konjugiert komplexen sind:

$$\left.\begin{aligned} \sigma^{k}{}_{\lambda\dot{\mu}}\,\psi^{\lambda}{}_{|k} &= \bar{a}\,\chi_{\dot{\mu}} \\ \sigma^{k\lambda\dot{\mu}}\,\chi_{\lambda|k} &= -\bar{a}\,\psi^{\dot{\mu}} \end{aligned}\right\}. \tag{46b}$$

Die Gleichungen (46) bilden die allgemein-relativistische Verallgemeinerung der Diracschen Gleichungen. Man sieht zunächst, daß sie eine allgemein-kovariante Form besitzen, und zwar sowohl gegenüber beliebigen Transformationen im Riemannschen Raume als gegenüber beliebigen Spintransformationen. Der Faktor a kann nach der bisherigen Herleitung irgendeine Ortsfunktion sein; da aber in (43) die Phasen von ψ und χ noch ganz willkürlich sind, kann man sie stets so einrichten, daß a rein imaginär wird. Da wir keinen Grund haben, dem Elektron eine von Punkt zu Punkt veränderliche Masse zuzuschreiben, nehmen wir a als eine Konstante an.

Wird das Koordinatensystem so gewählt, daß es im Punkte P_0 vollständig geodätisch ist (was nach § 5 immer möglich ist), so erhält man in der unmittelbaren Umgebung von P_0 die Diracschen Gleichungen in der gewohnten Form, wenn man für a und Φ_s

$$a = \frac{2\pi imc}{h\sqrt{2}}, \quad \Phi_s = \frac{2\pi}{h}\varphi_s$$

einsetzt, wo φ_s das elektromagnetische Potential des fremden Feldes ist. Die Welttensoren Φ_s, F_{pq}, welche die Parallelverschiebung und die Krümmung im Spinraum bestimmen, sollen daher (abgesehen von konstanten Faktoren) mit dem elektromagnetischen Potential und der Feldstärke identifiziert werden.

Die Gleichungen (46) definieren ein lineares selbstadjungiertes Eigenwertproblem für die 4 Komponenten ψ^{λ}, $\chi_{\dot{\lambda}}$. Um die Linearität einzusehen, hat man etwa die zweite Gleichung (46a) mit der ersten Gleichung (46b) zusammenzunehmen:

$$\left.\begin{aligned} \sigma^{k\mu\dot{\lambda}}\,\chi_{\dot{\lambda}|k} &= -a\,\psi^{\mu} \\ \sigma^{k}{}_{\dot{\mu}\lambda}\,\psi^{\lambda}{}_{|k} &= \bar{a}\,\chi_{\dot{\mu}} = -a\,\chi_{\dot{\mu}} \end{aligned}\right\}. \tag{47}$$

Führt man die Matrices

$$\left.\begin{aligned} {}'\sigma^{k} &= (\sigma^{k\alpha\dot{\beta}}), & \sigma^{k} &= (\sigma^{k}{}_{\dot{\alpha}\beta}), \\ \psi &= \begin{pmatrix}\psi^{1}\\ \psi^{2}\end{pmatrix}, & \chi &= \begin{pmatrix}\chi_{1}\\ \chi_{2}\end{pmatrix} \end{aligned}\right\} \tag{48}$$

ein, so kann man statt (47) auch schreiben:

$$\left.\begin{aligned} {}'\sigma^{k}\,\chi_{|k} &= -a\,\psi \\ \sigma^{k}\,\psi_{|k} &= -a\,\chi \end{aligned}\right\}$$

oder nach Multiplikation mit $\frac{h\sqrt{2}}{2\pi i}$

$$\left.\begin{aligned}\sqrt{2}\,{}'\sigma^k p_k \chi &= -mc\psi \\ \sqrt{2}\;\,\sigma^k p_k \psi &= -mc\psi\end{aligned}\right\}, \tag{49}$$

wo p_k der mit $\frac{h}{2\pi i}$ multiplizierte Operator der kovarianten Differentiation ist.

Um die Selbstadjungiertheit des Problems einzusehen, schreibe man (49) in 4reihige Matrices um:

$$\sum_k p_k \begin{pmatrix}\sqrt{2}\,\sigma^k & 0 \\ 0 & \sqrt{2}\,{}'\sigma^k\end{pmatrix}\begin{pmatrix}\psi \\ \chi\end{pmatrix} + \begin{pmatrix}0 & mc \\ mc & 0\end{pmatrix}\begin{pmatrix}\psi \\ \chi\end{pmatrix} = 0. \tag{50a}$$

Die hier auftretenden Matrices haben alle die Eigenschaft der Hermiteschen Symmetrie. Wählt man nun das Koordinatensystem im Spinraum so, daß σ^4 gleich der Einheitsmatrix wird (was immer möglich ist, da σ^4 eine positiv-definite Hermitesche Form definiert, welche immer in die Einheitsform transformiert werden kann), so hat die Gleichung (50) die Form

$$\left(\frac{h}{2\pi i}\frac{\partial}{\partial t} + H\right)\Psi = 0$$

mit selbstadjungiertem H und $\Psi = \begin{pmatrix}\psi \\ \chi\end{pmatrix}$.

Zum Vergleich mit den vierreihigen Matrixtheorien eignet sich eine andere Matrixform der Gleichungen (49) besser, nämlich

$$\sum p_k \begin{pmatrix}0 & \sqrt{2}\,{}'\sigma^k \\ \sqrt{2}\,\sigma^k & 0\end{pmatrix}\begin{pmatrix}\psi \\ \chi\end{pmatrix} = -mc\begin{pmatrix}\psi \\ \chi\end{pmatrix}. \tag{50b}$$

Die auf der linken Seite auftretenden Matrices sind die γ^k von Tetrode, welche die bekannten Relationen

$$\gamma^k\gamma^l + \gamma^l\gamma^k = 2g^{kl}$$

erfüllen: diese kommen nämlich auf

$$\left.\begin{aligned}{}'\sigma^k\,\sigma^l + {}'\sigma^l\,\sigma^k &= g^{kl} \\ \sigma^k\,{}'\sigma^l + \sigma^l\,{}'\sigma^k &= g^{kl}\end{aligned}\right\},$$

also auf (15) hinaus. Diese Matrices sind nicht Hermitesch; sie können es aber gemacht werden, indem man sie entweder (nach Schrödinger) mit γ_0 multipliziert und die zulässigen Koordinatensysteme durch geeignete

Vorschriften beschränkt oder (nach V. Bargmann[1]) durch Multiplikation mit einer anderen Matrix α die Hermitizität unabhängig vom Koordinatensystem erzwingt. Letzteres kommt im wesentlichen darauf hinaus, daß man zu der anderen Matrixgestalt (50a) übergeht, wo (wie wir sahen) alle Hermitizitätsschwierigkeiten verschwinden. Unsere Theorie unterscheidet sich also von den vierreihigen Matrixtheorien durch eine besondere Wahl der Koordinaten im vierdimensionalen Spinraum, wobei die Matrices γ^k die Gestalt (50b) annehmen und die vier Komponenten von Ψ sich in zwei Paare ψ^λ, $\chi_{\dot\lambda}$ trennen. Diese besondere Wahl ist aber in invarianter Weise dadurch ausgezeichnet, daß die beiden Paare bei Lorentztransformationen einzeln in sich transformiert werden.

Im ε-Formalismus ist

$$\varkappa^{\dot\lambda\mu} = \psi^{\dot\lambda}\psi^\mu + \chi^{\dot\lambda}\chi^\mu$$

kein Spinor, sondern eine Spinordichte vom Gewichte $+1$, da nur in diesem Falle eine eineindeutige Beziehung zwischen $\varkappa^{\dot\lambda\mu}$ und dem Weltvektor $\mathfrak{J}^s$ besteht. Die ψ^λ, χ^λ sind dann Spinordichten vom Absolutgewichte $+\frac{1}{2}$, ψ_λ, χ_λ, vom Absolutgewichte $-\frac{1}{2}$. Bei Ausführung der kovarianten Differentiation und in dem Krümmungstensor treten nur die $\overset{*}{\Gamma}{}^{\rho}_{\lambda s}$ auf. Sonst gelten alle Resultate dieses und der beiden folgenden Paragraphen auch im ε-Formalismus.

§ 8. Der Übergang zu den Gleichungen zweiter Ordnung.

Setzt man $\chi_{\dot\lambda}$ aus (46b) in (46a) ein, so erhält man:

$$\sigma^{k\dot\lambda\mu}\,\sigma^{l}_{\dot\lambda\rho}\,\psi^{\rho}{}_{|lk} = -\alpha\,\bar\alpha\,\psi^{\mu}. \tag{51}$$

Es soll hier die linke Seite dieser Gleichung ausgerechnet werden. Durch einfache Umschreibung und Anwendung von (39a) erhält man:

$$\left.\begin{aligned} -\alpha\,\bar\alpha\,\psi^{\mu} &= \tfrac{1}{2}\,\sigma^{k\dot\lambda\mu}\,\sigma^{l}_{\dot\lambda\rho}\,\psi^{\rho}{}_{|lk} + \tfrac{1}{2}\,\sigma^{l\dot\lambda\mu}\sigma^{k}_{\dot\lambda\rho}\,\psi^{\rho}{}_{|kl} \\ &= \tfrac{1}{2}\left(\sigma^{k\dot\lambda\mu}\,\sigma^{l}_{\dot\lambda\rho} + \sigma^{l\dot\lambda\mu}\sigma^{k}_{\dot\lambda\rho}\right)\psi^{\rho}{}_{|lk} + \tfrac{1}{2}\,\sigma^{l\dot\lambda\mu}\,\sigma^{k}_{\dot\lambda\rho}\,\psi^{\sigma}\,P^{\rho}_{\sigma lk} \end{aligned}\right\}. \tag{52}$$

Das erste Glied rechts ergibt nach (15)

$$\tfrac{1}{2}g^{kl}\,\psi^{\mu}{}_{|lk}\,.$$

Wir gehen zur Berechnung des zweiten Ausdruckes über. Es folgt aus (41):

$$\left.\begin{aligned} \tfrac{1}{2}\sigma^{l\dot\lambda\mu}\,\sigma^{k}_{\dot\lambda\rho}\,\psi^{\sigma}\,P^{\rho}_{\sigma lk} = \tfrac{1}{4}\,\sigma^{l\dot\lambda\mu}\sigma^{k}_{\dot\lambda\rho}\,\psi^{\sigma}\,R_{pslk}\,\sigma^{p\rho\dot\nu}\,\sigma^{s}_{\dot\nu\sigma} \\ + \frac{i}{4}\,\sigma^{l\dot\lambda\mu}\,\sigma^{k}_{\dot\lambda\rho}\,\psi^{\rho}\,F_{lk} \end{aligned}\right\}. \tag{53}$$

---

[1] V. Bargmann, S.-B. preuß. Ak. Wiss. 1932, S. 346.

Wir wollen zum Schluß beweisen:

$$\tfrac{1}{4} R_{pslk}\, \sigma^{p\rho\dot{\nu}}\, \sigma^{s}_{\dot{\nu}\sigma}\, \sigma^{l\dot{\lambda}\mu}\, \sigma^{k}_{\dot{\lambda}\rho}\, \psi^{\sigma} = \tfrac{1}{8} R\, \psi^{\mu}. \tag{54}$$

(R = Krümmungsskalar)

Die Umschreibung der linken Seite dieser Gleichung ergibt:

$$\left.\begin{aligned} \tfrac{1}{4} R_{pslk}\, \sigma^{l\dot{\lambda}\mu}\, \sigma^{k}_{\dot{\lambda}\rho}\, \sigma^{p\rho\dot{\nu}}\, \sigma^{s}_{\dot{\nu}\rho} = {} & \tfrac{1}{4} R_{pslk}\, \sigma^{l\dot{\lambda}\mu}\, \sigma^{k}_{\dot{\lambda}\rho}\, \sigma^{p\rho\dot{\nu}}\, \sigma^{s}_{\dot{\nu}\sigma} \\ & + \tfrac{1}{4} R_{pslk}\, \sigma^{l\dot{\lambda}\mu}\, \sigma^{p}_{\dot{\lambda}\rho}\, \sigma^{k\rho\dot{\nu}}\, \sigma^{s}_{\dot{\nu}\sigma} \\ & - \tfrac{1}{4} R_{kslp}\, \sigma^{l\dot{\lambda}\mu}\, \sigma^{k}_{\dot{\lambda}\rho}\, \sigma^{p\rho\dot{\nu}}\, \sigma^{s}_{\dot{\nu}\sigma} \end{aligned}\right\}. \tag{55}$$

Die zwei ersten Glieder ergeben in der Tat $\frac{1}{8} R\, \delta^{\mu}_{\sigma}$. Man ersieht dies, indem man (15) anwendet und die bekannten Definitionen von R_{kl}, R, berücksichtigt. Der Beweis von (54) wäre daher erbracht, wenn man zeigen könnte, daß das letzte Glied in (55) verschwindet.

Wir schreiben das letzte Glied in der Form:

$$-\tfrac{1}{12}\, \sigma^{l\dot{\lambda}\mu}\, \sigma^{k}_{\dot{\lambda}\rho}\, \sigma^{p\rho\dot{\nu}}\, \sigma^{s}{}_{\dot{\nu}\sigma} (R_{kslp} + R_{kslp} + R_{kslp}). \tag{56}$$

Wendet man jetzt wiederholt eine zu (55) ganz analoge Umformung, so erhält man aus (56)

$$-\tfrac{1}{12}\, \sigma^{l\dot{\lambda}\mu}\, \sigma^{k}_{\dot{\lambda}\rho}\, \sigma^{p\rho\dot{\nu}}\, \sigma^{s}_{\dot{\nu}\sigma} [R_{kslp} - R_{pslk} - R_{lskp}].$$

Infolge der Symmetrieeigenschaften des Krümmungstensors kann der Klammerausdruck auch in der Form

$$R_{kslp} + R_{klps} + R_{kpsl}$$

geschrieben werden. Das verschwindet aber identisch und daher ist auch (54) bewiesen.

Die Gleichung (51), die den Ausgangspunkt dieser Rechnungen bildete, kann daher in der Form

$$g^{lk}\, \psi^{\mu}{}_{|lk} + \tfrac{1}{4} R\, \psi^{\mu} + \frac{i}{2} F_{lk}\, \sigma^{l\dot{\lambda}\mu}\, \sigma^{k}_{\dot{\lambda}\sigma}\, \psi^{\sigma} = -2\, \alpha \bar{\alpha}\, \psi^{\mu} \tag{57a}$$

geschrieben werden. Analog findet man für $\chi_{\dot{\mu}}$:

$$g^{lk}\, \chi_{\dot{\mu}|lk} + \tfrac{1}{4} R\, \chi_{\dot{\mu}} + \frac{i}{2} F_{lk}\, \sigma^{l}{}_{\lambda\dot{\mu}}\, \sigma^{k\lambda\dot{\sigma}}\, \chi_{\dot{\sigma}} = -2\, \alpha \bar{\alpha}\, \chi_{\dot{\mu}}. \tag{57b}$$

Das erste Glied auf der linken Seite und die rechte Seite von (57) entsprechen den in der Gordon-Kleinschen Gleichung auftretenden Ausdrücken. Außer diesen und außer dem (die elektromagnetische Feldstärke enthaltendem) Spingliede tritt in diesen Gleichungen noch der Krümmungsskalar auf.

§ 9. Der Energie-Impuls-Tensor.

Kann aus den das Wellenfeld charakterisierenden Größen der Energie-Impuls-Tensor gebildet werden? Er müßte reell sein und den folgenden Gleichungen genügen:

$$\left.\begin{aligned} T^{kl} &= T^{lk} \\ T^{lk}{}_{|l} &= F^{sk}\mathfrak{J}_s \end{aligned}\right\}. \tag{58}$$

Man kann zeigen, daß der zweiten dieser Bedingungen der Welttensor

$$\left.\begin{aligned} 'T^l_k &= i(\psi^\lambda \sigma^l_{\lambda\dot\mu}\psi^{\dot\mu}{}_{|k} - \psi^{\dot\lambda}\sigma^l{}_{\dot\lambda\mu}\psi^\mu{}_{|k}) \\ &- i(\chi_\lambda \sigma^{l\lambda\dot\mu}\chi_{\dot\mu|k} - \chi_{\dot\lambda}\sigma^{l\dot\lambda\mu}\chi_{\mu|k}) \end{aligned}\right\} \tag{59}$$

genügt. Man erkennt gleich, daß $'T^l_k$ reell ist. Zwecks Berechnung von $'T^l_{k|l}$ wollen wir zunächst feststellen, welchen Beitrag zu $'T^l{}_{k|l}$ das erste Glied von (59) liefert. Es ist:

$$\left.\begin{aligned} i(\psi^\lambda \sigma^l_{\lambda\dot\mu}\psi^{\dot\mu}{}_{|k})_{|l} &= i\bar a\chi_{\dot\mu}\psi^{\dot\mu}{}_{|k} + i a\psi^\mu \chi_{\mu|k} \\ &+ \frac{i}{2}\sigma^l_{\lambda\dot\mu}R_{prlk}\sigma^{p\dot\mu\nu}\sigma^r_{\dot\mu\nu}\psi^\lambda\psi^{\dot\mu} \\ &+ \tfrac{1}{2}F_{lk}\sigma^l_{\lambda\dot\mu}\psi^\lambda\psi^{\dot\mu} \end{aligned}\right\}. \tag{60}$$

Wird das den Krümmungstensor enthaltende Glied auf ganz analoge Weise wie im vorigen § umgeformt, so erhält man:

$$\left.\begin{aligned} i(\psi^\lambda \sigma^l_{\lambda\dot\mu}\psi^{\dot\mu}{}_{|k})_{|l} &= i\bar a\chi_{\dot\mu}\psi^{\dot\mu}{}_{|k} + i a\psi^\mu \chi_{\mu|k} \\ &- \frac{i}{2}R_{kl}\sigma^l_{\lambda\dot\mu}\psi^\lambda\psi^{\dot\mu} + \tfrac{1}{2}F_{lk}\sigma^l_{\lambda\dot\mu}\psi^\lambda\psi^{\dot\mu} \end{aligned}\right\}. \tag{61}$$

Die Berechnung der anderen Glieder und Summation ergibt:

$$'T^{lk}{}_{|l} = F^{sk}\mathfrak{J}_s. \tag{62}$$

Der Tensor $'T^{lk}$ ist reell, genügt dem Erhaltungssatze, ist aber nicht symmetrisch. Kann der Tensor $'T^{lk}$ symmetrisiert werden? D. h.: Genügt auch der Tensor

$$T^{lk} = \tfrac{1}{2}('T^{lk} + 'T^{kl}) \tag{63}$$

dem Erhaltungssatze? Es soll gezeigt werden, daß es in der Tat der Fall ist. Wir wollen daher jetzt nicht, wie vorher $'T^{lk}{}_{|l}$, sondern $'T^{kl}{}_{|l}$ berechnen.

Für $'T^{kl}$ haben wir den folgenden Ausdruck:

$$\left.\begin{aligned} 'T^{kl} &= i g^{sl}(\psi^\lambda \sigma^k_{\lambda\dot\mu}\psi^{\dot\mu}{}_{|s} - \psi^{\dot\lambda}\sigma^k_{\dot\lambda\mu}\psi^\mu{}_{|s}) \\ &- i g^{sl}(\chi_\lambda \sigma^{k\lambda\dot\mu}\chi_{\dot\mu|s} - \chi_{\dot\lambda}\sigma^{k\dot\lambda\mu}\chi_{\mu|s}) \end{aligned}\right\}. \tag{64}$$

Wir berechnen auch jetzt den vom ersten Gliede gelieferten Beitrag zu $'T^{kl}{}_{|l}$, wobei wir die Gleichungen (57) berücksichtigen:

$$\left.\begin{aligned} i g^{sl}(\psi^{\lambda}\sigma^{k}_{\lambda\dot{\mu}}\psi^{\dot{\mu}}{}_{|s})_{|l} &= i g^{sl}\psi^{\lambda}{}_{|l}\sigma^{k}_{\lambda\dot{\mu}}\psi^{\dot{\mu}}{}_{|s} + i g^{sl}\psi^{\lambda}\sigma^{k}_{\lambda\dot{\mu}}\psi^{\dot{\mu}}{}_{|sl} \\ &= i g^{sl}\psi^{\lambda}{}_{|l}\sigma^{k}_{\lambda\dot{\mu}}\psi^{\dot{\mu}}{}_{|s} - 2\alpha\bar{\alpha} i\psi^{\dot{\mu}}\psi^{\lambda}\sigma^{k}_{\lambda\dot{\mu}} \\ &\quad - \frac{i}{4}R\psi^{\dot{\mu}}\psi^{\lambda}\sigma^{k}_{\lambda\dot{\mu}} - \tfrac{1}{2}F_{pr}\sigma^{p\rho\dot{\mu}}\sigma^{r}_{\rho\dot{\sigma}}\psi^{\dot{\sigma}}\psi^{\lambda}\sigma^{k}_{\lambda\dot{\mu}} \end{aligned}\right\}. \tag{65}$$

Nach Ausführung einer analogen Rechnung an den anderen Gliedern und Summation erhält man:

$$'T^{kl}{}_{|l} = F^{sk}\mathfrak{J}_{s}. \tag{66}$$

Daraus folgt aber, daß der durch (63) definierte Tensor symmetrisch ist und den Erhaltungssätzen genügt, d. h. die Bedingungen (58) erfüllt. Er gibt den Beitrag des einen betrachteten Elektrons zum Energie-Impulstensor der Welt.

Wir wollen jetzt den einfachsten Fall annehmen, und zwar den des freien Elektrons. Es verschwindet dann der Tensor F_{sk}, und wir haben einfach

$$T^{lk}{}_{|k} = 0. \tag{67}$$

Die Kenntnis des Energie-Impuls-Tensors ermöglicht die Aufstellung der Gravitationsgleichungen. Für den Fall des freien Elektrons und beim Fehlen von fremden Massen im betrachteten Gebiet ist:

$$R^{l}_{k} - \tfrac{1}{2}\delta^{l}_{k}R = \varkappa T^{l}_{k} \tag{68}$$

($\varkappa$ = Gravitationskonstante). Die Verjüngung in den Indices k und l ergibt

$$-R = \varkappa T = 2i\varkappa(\alpha\chi_{\lambda}\psi^{\lambda} - \bar{\alpha}\chi_{\dot{\lambda}}\psi^{\dot{\lambda}}). \tag{69}$$

Der Skalar $T = T^{k}_{k}$ kann daher (bei entsprechender Normierung von ψ^{μ} und $\chi_{\dot{\mu}}$) mit der Massendichte identifiziert werden.

Wenn es erlaubt ist, die Gravitationswirkung des freien Elektrons auf sich selbst dadurch zu berechnen, daß man den aus (69) sich ergebenden Wert von R in (57) einsetzt, so ergeben sich die folgenden nichtlinearen Gleichungen:

$$\left.\begin{aligned} g^{kl}\psi^{\mu}{}_{|lk} - \frac{i}{2}\varkappa\psi^{\mu}(\alpha\chi_{\lambda}\psi^{\lambda} - \bar{\alpha}\chi_{\dot{\lambda}}\psi^{\dot{\lambda}}) &= -2\alpha\bar{\alpha}\psi^{\mu} \\ g^{kl}\psi_{\dot{\mu}|lk} - \frac{i}{2}\varkappa\chi_{\dot{\mu}}(\alpha\chi_{\lambda}\psi^{\lambda} - \bar{\alpha}\chi_{\dot{\lambda}}\psi^{\dot{\lambda}}) &= -2\alpha\bar{\alpha}\chi_{\dot{\mu}} \end{aligned}\right\}. \tag{70}$$

Das mit $\varkappa$ behaftete Glied ist selbstverständlich sehr klein im Vergleich mit $2\,\alpha\bar{\alpha}$, so daß man in erster Annäherung in einem in P_0 vollständig geodätischen System die Gordon-Kleinsche Gleichung erhält.

Ausgegeben am 30. Mai.

Berichtigung

zu der Arbeit L. Infeld und B. L. van der Waerden, Die Wellengleichung des Elektrons in der allgemeinen Relativitätstheorie.

Hr. Prof. J. A. Schouten macht uns in freundlicher Weise auf folgende Fehler aufmerksam.

In § 4, letzter Absatz (kleine Typen) soll statt »Gewicht $\frac{1}{2}$« und »Gewicht $-\frac{1}{2}$« stehen »Absolutgewicht $\frac{1}{2}$« und »Absolutgewicht $-\frac{1}{2}$«. Desgleichen in § 5, letzter Absatz.

Der letzte Absatz von § 7 (kleine Typen) soll so heißen:

»Im ε-Formalismus ist

$$\varkappa^{\dot{\lambda}\mu} = \psi^{\dot{\lambda}}\psi^{\mu} + \chi^{\dot{\lambda}}\chi^{\mu} = \psi^{\dot{\lambda}}\psi^{\mu} + \varepsilon^{\dot{\lambda}\dot{\rho}}\varepsilon^{\mu\sigma}\chi_{\dot{\rho}}\chi_{\sigma} \tag{50c}$$

kein Spinor, sondern eine Spinordichte vom Absolutgewicht $+1$, da nur in diesem Falle eine eineindeutige Beziehung zwischen $\varkappa^{\dot{\lambda}\mu}$ und dem Weltvektor J^s besteht. Es liegt also nahe, unter ψ^{λ} (und $\psi^{\dot{\lambda}}$) Spinoren vom Absolutgewicht $\frac{1}{2}$ zu verstehen. Die Diracschen Gleichungen (47) verlangen dann, daß χ_{μ} und $\chi_{\dot{\mu}}$ Spinoren vom Absolutgewicht $-\frac{1}{2}$ seien. In den Transformationsformeln für χ^{μ} und ψ_{λ} gehen dann aber die Zahlenfaktoren $|\Delta|^{\frac{1}{2}}\Delta^{-1}$ bzw. $|\Delta|^{-\frac{1}{2}}\Delta$ ein. Wenn wir also verabreden, uns die Spinordichten ψ^{λ} und χ_{μ} vom Absolutgewicht $\pm\frac{1}{2}$, die in (47) auch allein vorkommen, als physikalisch-bedeutsame Größen zu betrachten, die ψ_{λ} und χ^{μ} dagegen als reine Rechengrößen, die aus den endgültigen Formeln immer eliminiert werden können (wie in 50c), so gilt für die physikalisch-bedeutsamen Größen der Ausspruch von § 4, daß sie alle Vektordichten vom Gewicht $\pm\frac{1}{2}$ (+ für kontra, — für kovariant) sind oder sich aus solchen Dichten zusammensetzen. In den Formeln für ihre kovariante Differentiation treten nur die $\overset{*}{\Gamma}{}^{\rho}_{\lambda s}$ auf.«

Ausgegeben am 3. Juli.

Berlin, gedruckt in der Reichsdruckerei.

Chapter 15
Albert Einstein, Leopold Infeld and Banesh Hoffmann (1938): The Gravitational Equations and the Problem of Motion

Albert Einstein, Leopold Infeld and Banesh Hoffmann (1938). The Gravitational Equations and the Problem of Motion. *The Annals of Mathematics*, 39: 65–100.

ANNALS OF MATHEMATICS
Vol. 39, No. 1, January, 1938

THE GRAVITATIONAL EQUATIONS AND THE PROBLEM OF MOTION

BY A. EINSTEIN, L. INFELD, AND B. HOFFMANN

(Received June 16, 1937)

Introduction. In this paper we investigate the fundamentally simple question of the extent to which the relativistic equations of gravitation determine the motion of ponderable bodies.

Previous attacks on this problem[1] have been based upon gravitational equations in which some specific energy-momentum tensor for matter has been assumed. Such energy-momentum tensors, however, must be regarded as purely temporary and more or less phenomenological devices for representing the structure of matter, and their entry into the equations makes it impossible to determine how far the results obtained are independent of the particular assumption made concerning the constitution of matter.

Actually, the only equations of gravitation which follow without ambiguity from the fundamental assumptions of the general theory of relativity are the equations for empty space, and it is important to know whether they *alone* are capable of determining the motion of bodies. The answer to this question is not at all obvious. It is possible to find examples in classical physics leading to either answer, yes or no. For instance, in the ordinary Maxwell equations for empty space, in which electrical particles are regarded as point singularities of the field, the motion of these singularities is not determined by the linear field equations. On the other hand, the well-known theory of Helmholtz on the motion of vortices in a non-viscous fluid gives an instance where the motion of line singularities is actually determined by partial differential equations alone, which are there non-linear.

We shall show in this paper that the gravitational equations for empty space are in fact sufficient to determine the motion of matter represented as point singularities of the field. The gravitational equations are non-linear, and, because of the necessary freedom of choice of the coördinate system, are such that four differential relations exist between them so that they form an overdetermined system of equations. The overdetermination is responsible for the existence of equations of motion, and the non linear character for the existence of terms expressing the interaction of moving bodies.

Two essential steps lead to the determination of the motion.

---

[1] Droste, Ac. van Wet. Amsterdam *19*, 447 (1916). De Sitter, Monthly Notices of the R. A. S. *67*, 155 (1916). Mathisson, Zeits. f. Physik, *67*, 270, 826 (1931), *69*, 389 (1931). Levi-Civita, Am. Jour. of Math., lix, 3, 225 (1937).

(1) By means of a new method of approximation, specially suited to the treatment of quasi-stationary fields, the gravitational field due to moving particles is determined.
(2) It is shown that for two-dimensional spatial surfaces containing singularities certain surface integral conditions are valid which determine the motion.

In the second part of this paper we actually calculate the first two non-trivial stages of the approximation. In the first of these the equations of motion take the Newtonian form. In the second the equations of motion, which we calculate only for the case of two massive particles, take a more complicated form but do not involve third or higher derivatives with respect to the time.

The method is, in principle, applicable for any order of approximation, the problem reducing to specific integrations at each stage, but we have not proved that higher time derivatives than the second will not ultimately occur in the equations of motion.

In the determination of the field and the equations of motion non-Galilean values at infinity and singularities of the type of dipoles, quadrupoles, and higher poles, must be excluded from the field in order that the solution shall be unique.

It is of significance that our equations of motion do not restrict the motion of the singularities more strongly than the Newtonian equations, but this may be due to our simplifying assumption that matter is represented by singularities, and it is possible that it would not be the case if we could represent matter in terms of a field theory from which singularities were excluded. The representation of matter by means of singularities does not enable the field equations to fix the sign of mass so that, so far as the present theory is concerned, it is only by convention that the interaction between two bodies is always an attraction and not a repulsion. A possible clue as to why the mass must be positive can be expected only from a theory which gives a representation of matter free from singularities.[2]

Our method can be applied to the case when the Maxwell energy-momentum tensor is included in the field equations and, as is shown in part II, it leads to a derivation of the Lorentz force.

In the Maxwell-Lorentz electrodynamics, as also in the earlier approximation method for the solution of the gravitational equations, the problem of determining the field due to moving bodies is solved through the integration of wave equations by retarded potentials. The sign of the flow of time there plays a decisive rôle since, in a certain sense, the field is expanded in terms of only those waves which proceed towards infinity. In our theory, however, the equations to be solved at each stage of the approximation are not wave equations but merely spatial potential equations. Since such equations as those of the gravitational and of the electromagnetic field are actually invariant under a

[2] Einstein and Rosen, Phys. Rev. vol. *48*, 73 (1935).

reversal of the sign of time, it would seem that the method presented here, is the natural one for their solution. Our method, in which the time direction is not distinguished, corresponds to the introduction of standing waves in the wave equation and cannot lead to the conclusion that in the circular motion of two point masses energy is radiated to infinity in the form of waves.

I. General Theory

1. **Field Equations and Coördinate Conditions.** Since it is an essential part of the work to make a separation between space and time we shall, throughout this paper, use the convention that Latin indices take on only the spatial values 1, 2, 3 while Greek indices refer to both space and time, running over the values 0, 1, 2, 3.

As explained in the introduction, we discuss only the gravitational equations for empty space, treating the sources of the field as singularities. If we denote the ordinary derivative of a quantity by means of a line followed by the appropriate suffix, as

$$(1,1) \qquad \frac{\partial g_{\mu\nu}}{\partial x^\sigma} \to g_{\mu\nu|\sigma}\,; \qquad \frac{\partial^2 g_{\mu\nu}}{\partial x^\sigma \partial x^\rho} \to g_{\mu\nu|\sigma\rho}\,,$$

we may write the field equations in the form

$$(1,2) \qquad R_{\mu\nu} = -\left\{ \begin{matrix} \lambda \\ \mu\nu \end{matrix} \right\}_{|\lambda} + \left\{ \begin{matrix} \lambda \\ \mu\lambda \end{matrix} \right\}_{|\nu} + \left\{ \begin{matrix} \lambda \\ \mu\sigma \end{matrix} \right\}\left\{ \begin{matrix} \sigma \\ \lambda\nu \end{matrix} \right\} - \left\{ \begin{matrix} \lambda \\ \mu\nu \end{matrix} \right\}\left\{ \begin{matrix} \sigma \\ \lambda\sigma \end{matrix} \right\} = 0.$$

Let the symbols $\eta_{\mu\nu}$, $\eta^{\mu\nu}$ be defined by

$$(1,3) \qquad \eta_{\mu\nu} = \eta^{\mu\nu} = \begin{Bmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{Bmatrix},$$

so that they represent the metric of empty space-time. Then if we introduce the quantities $h_{\mu\nu}$, $h^{\mu\nu}$ by the relations

$$(1,4) \qquad g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}\,, \qquad g^{\mu\nu} = \eta^{\mu\nu} + h^{\mu\nu},$$

the $h_{\mu\nu}$ and $h^{\mu\nu}$ will represent the deviation of space-time from the flat case. The $h^{\mu\nu}$ can be calculated as functions of the $h_{\mu\nu}$ by means of the relations

$$(1,5) \qquad g_{\mu\nu} g^{\nu\sigma} = \delta_\mu^\sigma\,.$$

In general the $h_{\mu\nu}$ will be small relative to unity, but we make no assumptions here concerning their order of magnitude.

By means of (1, 4), (1, 5) we can express the components of $R_{\mu\nu}$ as functions of the $h_{\mu\nu}$, and for reasons which will become clear when we come to the method of approximation used in the present work we separate the various terms so obtained into two groups in the following manner. First we separate the terms linear in the h's from those which are quadratic and of higher order. At this stage of the separation the field equations are of the form

$$R_{00} = \tfrac{1}{2}\{-h_{00|ss} + 2h_{0s|0s} - h_{ss|00}\} + L'_{00} = 0, \tag{1, 6}$$

$$R_{0n} = \tfrac{1}{2}\{-h_{0n|ss} + h_{0s|ns} + h_{ns|0s} - h_{ss|n0}\} + L'_{0n} = 0, \tag{1, 7}$$

$$R_{mn} = \tfrac{1}{2}\{-h_{mn|ss} + h_{ms|ns} + h_{ns|ms} - h_{ss|mn} + h_{mn|00} - h_{m0|n0} - h_{n0|m0} + h_{00|mn}\} + L'_{mn} = 0, \tag{1, 8}$$

where the $L'_{\mu\nu}$ represent the non-linear terms. We now take

(1, 9) from R_{00} the terms $h_{0s|0s} - \tfrac{1}{2}h_{ss|00}$,

(1, 10) from R_{0n} no terms

(1, 11) and from R_{mn} the terms $-\tfrac{1}{2}h_{0m|0n} - \tfrac{1}{2}h_{0n|0m} + \tfrac{1}{2}h_{mn|00}$,

and add them to the non-linear group. Introducing the symbol $L_{\mu\nu}$ to denote the non-linear group $L'_{\mu\nu}$ together with these added linear members, we may write the field equations in the separated from

$$R_{00} = -\tfrac{1}{2}h_{00|ss} + L_{00} = 0, \tag{1, 12}$$

$$R_{0n} = -\tfrac{1}{2}h_{0n|ss} + \tfrac{1}{2}(h_{ns} - \tfrac{1}{2}\delta_{ns}h_{ll} + \tfrac{1}{2}\delta_{ns}h_{00})_{|0s} - \tfrac{1}{4}(h_{00} + h_{ss})_{|0n} + \tfrac{1}{2}h_{0s|ns} + L_{0n} = 0, \tag{1, 13}$$

$$R_{mn} = -\tfrac{1}{2}h_{mn|ss} + \tfrac{1}{2}(h_{ms} - \tfrac{1}{2}\delta_{ms}h_{ll} + \tfrac{1}{2}\delta_{ms}h_{00})_{|ns} + \tfrac{1}{2}(h_{ns} - \tfrac{1}{2}\delta_{ns}h_{ll} + \tfrac{1}{2}\delta_{ns}h_{00})_{|ms} + L_{mn} = 0, \tag{1, 14}$$

where the L's are given explicitly by the formulas

$$L_{00} = h_{0s|0s} - \tfrac{1}{2}h_{ss|00} - (h^{\lambda\sigma}[00, \sigma])_{|\lambda} + (h^{\lambda\sigma}[0\lambda, \sigma])_{|0} + \begin{Bmatrix}\lambda\\0\sigma\end{Bmatrix}\begin{Bmatrix}\sigma\\\lambda 0\end{Bmatrix} - \begin{Bmatrix}\lambda\\00\end{Bmatrix}\begin{Bmatrix}\sigma\\\lambda\sigma\end{Bmatrix}, \tag{1, 15}$$

$$L_{0n} = -(h^{\lambda\sigma}[0n, \sigma])_{|\lambda} + (h^{\lambda\sigma}[n\lambda, \sigma])_{|0} + \begin{Bmatrix}\lambda\\n\sigma\end{Bmatrix}\begin{Bmatrix}\sigma\\\lambda 0\end{Bmatrix} - \begin{Bmatrix}\lambda\\n0\end{Bmatrix}\begin{Bmatrix}\sigma\\\lambda\sigma\end{Bmatrix}, \tag{1, 16}$$

$$L_{mn} = -\tfrac{1}{2}h_{0m|0n} - \tfrac{1}{2}h_{0n|0m} + \tfrac{1}{2}h_{mn|00} - (h^{\lambda\sigma}[mn, \sigma])_{|\lambda} + (h^{\lambda\sigma}[m\lambda, \sigma])_{|n} + \begin{Bmatrix}\lambda\\m\sigma\end{Bmatrix}\begin{Bmatrix}\sigma\\\lambda n\end{Bmatrix} - \begin{Bmatrix}\lambda\\mn\end{Bmatrix}\begin{Bmatrix}\sigma\\\lambda\sigma\end{Bmatrix}. \tag{1, 17}$$

If we introduce the quantities $\gamma_{\mu\nu}$ defined by

$$\gamma_{\mu\nu} = h_{\mu\nu} - \tfrac{1}{2}\,\eta_{\mu\nu}\eta^{\sigma\rho}h_{\sigma\rho}\,, \tag{1, 18}$$

or, in expanded form,

$$\gamma_{00} = \tfrac{1}{2}h_{00} + \tfrac{1}{2}h_{ll}\,, \tag{1, 19}$$

$$\gamma_{0n} = h_{0n}\,, \tag{1, 20}$$

$$\gamma_{mn} = h_{mn} - \tfrac{1}{2}\delta_{mn}h_{ll} + \tfrac{1}{2}\delta_{mn}h_{00}\,, \tag{1, 21}$$

we may write the field equations (1, 12), (1, 13), (1, 14) in the form

$$R_{00} = -\tfrac{1}{2}h_{00|ss} + L_{00} = 0, \tag{1, 22}$$

$$R_{0n} = -\tfrac{1}{2}h_{0n|ss} + \tfrac{1}{2}\gamma_{ns|0s} + \tfrac{1}{2}(\gamma_{0s|s} - \gamma_{00|0})_{|n} + L_{0n} = 0, \tag{1, 23}$$

$$R_{mn} = -\tfrac{1}{2}h_{mn|ss} + \tfrac{1}{2}\gamma_{ms|sn} + \tfrac{1}{2}\gamma_{ns|sm} + L_{mn} = 0. \tag{1, 24}$$

Since there are four identities between these field equations, we may impose four coördinate conditions, in the form of four non-tensorial equations involving the gravitational potentials, so as to limit the arbitrariness of the solutions by limiting the freedom of choice of the coördinate system. It turns out to be simplest to use coördinate conditions which involve only quantities which enter the explicitly written parts of the field equations (1, 23), (1, 24). These equations, in fact, suggest that we take as our coördinate conditions[3]

$$\gamma_{0s|s} - \gamma_{00|0} = 0, \tag{1, 25}$$

$$\gamma_{ms|s} = 0. \tag{1, 26}$$

With these coördinate conditions the field equations become merely

$$h_{00|ss} = 2L_{00}\,, \tag{1, 27}$$

$$h_{0n|ss} = 2L_{0n}\,, \tag{1, 28}$$

$$h_{mn|ss} = 2L_{mn}\,. \tag{1, 29}$$

For the further argument it is necessary that we write these equations in such a way that the Laplacians of the γ's enter instead of the Laplacians of the h's. We therefore replace the above equations by the equivalent equations

$$\gamma_{00,ss} = 2\Lambda_{00}\,, \tag{1, 30}$$

$$\gamma_{0n,ss} = 2\Lambda_{0n}\,, \tag{1, 31}$$

$$\gamma_{mn,ss} = 2\Lambda_{mn}\,, \tag{1, 32}$$

[3] The choice of the coördinate conditions is, to a large extent arbitrary, and it might seem rather more natural to use the conditions

$$\eta^{\nu\mu}\gamma_{\alpha\nu|\mu} = 0$$

which are invariant under a Lorentz transformation. However, it turns out that the actual calculation of the field is simpler when we use the coördinate conditions given in the text and it is for this reason that we employ it in the general theory.

where Λ is related to L exactly as γ is to h:

$$\Lambda_{00} = \tfrac{1}{2}L_{00} + \tfrac{1}{2}L_{ll}\,, \tag{1, 33}$$

$$\Lambda_{0n} = L_{0n}\,, \tag{1, 34}$$

$$\Lambda_{mn} = L_{mn} - \tfrac{1}{2}\delta_{mn}L_{ll} + \tfrac{1}{2}\delta_{mn}L_{00}\,. \tag{1, 35}$$

These field equations, (1, 33), (1, 34), (1, 35) together with the coördinate conditions (1, 25), (1, 26) will form the basis of our further considerations.

2. **Fundamental Integral Properties of the Field.** Let us consider three functions A_n; $(n = 1, 2, 3)$. They need not be tensors. From these functions we may build the three further functions

$$(A_{n|s} - A_{s|n})_{|s}\,, \tag{2, 1}$$

which can be explicitly written as

$$\begin{gathered}\{(A_{1|2} - A_{2|1})_{|2} - (A_{3|1} - A_{1|3})_{|3}\}, \quad \{(A_{2|3} - A_{3|2})_{|3} - (A_{1|2} - A_{2|1})_{|1}\}, \\ \{(A_{3|1} - A_{1|3})_{|1} - (A_{2|3} - A_{3|2})_{|2}\}.\end{gathered} \tag{2, 2}$$

These three functions thus constitute the curl of the three functions

$$(A_{2|3} - A_{3|2}), \qquad (A_{3|1} - A_{1|3}), \qquad (A_{1|2} - A_{2|1}). \tag{2, 3}$$

Consider any surface S which does not pass through singularities of the field. Since (2, 1) is the curl of (2, 3), it follows from Stokes' theorem that the integral of the "normal"[4] component of (2, 1) over S is equal to the line integral of the tangential component of (2, 3) taken around the rim of S. If S is a closed surface its rim is of zero length so that the latter integral will vanish. We therefore have the theorem that, if S is any closed surface which does not pass through singularities of the field, then

$$\int (A_{n|s} - A_{s|n})_{|s} \cos(\mathbf{n}\cdot\mathbf{N})\, dS = 0, \tag{2, 4}$$

where $(\mathbf{n}\cdot\mathbf{N})$ denotes the "angle" between the direction of x^n and the "normal" to S, and the summation convention applies to the n. This theorem is valid whether S encloses singularities or not, and we shall now apply it to the present problem.

[4] Words like *normal, angle, sphere,* and so on are used here in a purely conventional sense to designate the corresponding functions of the coördinates x^m and equations which are implied by these terms in Euclidean geometry. The argument of this paragraph is independent of any particular metric, and we use the Euclidean nomenclature merely because it is apt and convenient.

From the coördinate conditions (1, 25), (1, 26) and the field equations (1, 31), (1, 32) we have

(2, 5) $$(\gamma_{0n|s} - \gamma_{0s|n})_{|s} = 2\Lambda_{0n} - \gamma_{00|0n} ,$$

(2, 6) $$(\gamma_{mn|s} - \gamma_{ms|n})_{|s} = 2\Lambda_{mn} .$$

We see that the left-hand sides of (2, 5), (2, 6) give four quantities of the form (2, 1), one coming from (2, 5) and three from (2, 6) for $m = 1, 2, 3$. It follows from (2, 4) that, if S is a surface which does not pass through singularities of the field,

(2, 7) $$\int (\gamma_{00|0n} - 2\Lambda_{0n}) \cos (\mathbf{n}\cdot\mathbf{N})\, dS = 0,$$

(2, 8) $$\int 2\Lambda_{mn} \cos (\mathbf{n}\cdot\mathbf{N})\, dS = 0.$$

From (2, 5), (2, 6) we see that, in those regions where there are no singularities,

(2, 9) $$(\gamma_{00|0n} - 2\Lambda_{0n})_{|n} = 0,$$

(2, 10) $$(2\Lambda_{mn})_{|n} = 0.$$

Therefore Gauss' theorem shows that if we take two closed surfaces S, S' such that no singularity lies on or between S and S', the integrals over S and S' give the same result. But the validity of the integral conditions for surfaces which enclose singularities, or more generally, which enclose regions where the field equations for empty space are not fulfilled, can only be shown by means of Stokes' theorem.

We are treating matter as a singularity in the field. Let us assume there are p bodies, each represented by a point singularity. The coördinates of each such singularity will be functions of the time alone. Since (2, 7), (2, 8) are valid for any S provided only that it does not pass through a singularity, we may choose p such surfaces, each enclosing only one of the p singularities, and thus obtain $4p$ distinct integral conditions. Each of these, being now independent of the shape of its S, will give a relation between the coördinates of the singularities and their time derivatives, and we shall see later that the integral conditions give, in fact, the *equations of motion* of the singularities. These equations are derived here from the field equations and coördinate conditions alone without any extraneous assumption.

If, instead of integrating around one singularity at a time, we integrate over a surface which contains all the singularities, we obtain the laws of conservation of energy and linear momentum for the whole system. These laws are, of course, merely consequences of the laws of motion for the individual particles but owing to many cancellations they take a comparatively simple form.

3. **The Method of Approximation.** The method of approximation which has been used up to now in the theory of relativity is as follows. We consider that in the equation

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \tag{3, 1}$$

the $h_{\mu\nu}$ depend continuously on a positive parameter λ in such a way that they vanish for $\lambda = 0$, so that for $\lambda = 0$ space-time becomes Galilean. We assume, therefore, that the $h_{\mu\nu}$ can be expanded in a power series[5] in λ:

$$h_{\mu\nu} = \sum_{l=1}^{\infty} \lambda^l \underset{l}{h_{\mu\nu}}. \tag{3, 2}$$

This expansion is introduced in the field equations which are then grouped according to the different powers of λ, taking the form

$$0 = R_{\mu\nu} = \sum_{l=1}^{\infty} \lambda^l \underset{l}{R_{\mu\nu}}. \tag{3, 3}$$

In order that a set of $h_{\mu\nu}$ depending on the parameter λ shall exist as a solution of the field equations it is necessary that each of the equations

$$\underset{l}{R_{\mu\nu}} = 0 \tag{3, 4}$$

shall be satisfied. The best known example of this method is its application to the first approximation.

We shall now show why this method of approximation is unsuitable for the treatment of quasi-stationary fields. If we introduce an energy tensor for the matter which produces the field we obtain for the first approximation, using imaginary time, the well-known equations

$$\gamma_{\mu\nu|\sigma\sigma} = -2T_{\mu\nu}, \tag{3, 5}$$

where the coördinate system is determined by the equations

$$\gamma_{\mu\sigma|\sigma} = 0. \tag{3, 6}$$

In the simplest case of incoherent matter (dust) producing the field we have

$$T_{\mu\nu} = \rho \frac{d\xi^\mu}{ds}\frac{d\xi^\nu}{ds}, \tag{3, 7}$$

where $d\xi^\mu/ds$ are the components of the velocity measured in terms of the proper time s. If we are concerned with a quasi-static situation, $d\xi^0/ds$ is of the order of magnitude of unity while the $d\xi^m/ds$ are relatively small. Thus in such a case we shall have

$$| T_{00} | \gg | T_{0n} | \gg | T_{mn} |, \tag{3, 8}$$

[5] In λ^l the l will always be an exponent, not a contravariant index!

and from the equations (3, 5) we must have correspondingly

$$|\gamma_{00}| \gg |\gamma_{0n}| \gg |\gamma_{mn}| . \tag{3, 9}$$

The usual method of approximation does not take this into account since it treats all the γ's as of the same order of magnitude although, in the quasi-static case, γ_{00} is very much larger than the other components of $\gamma_{\mu\nu}$. A really good method of approximation for the quasi-static case should make essential use of the relations (3, 9).

We are led to our present method of approximation most simply by considering the problem of constructing a method of approximation which is suitable for the solution of the approximate field equations (3, 5) for the quasi-static case. It turns out that the method of approximation to which we are led in this way is also suitable for the solution of the rigorous gravitational equations even when we are not dealing with quasi-static cases.

The first step is to give an explicit expression for the fact that the time derivative of a field quantity is small relative to the quantity itself and to its spatial derivatives. To do this we introduce an auxiliary time coördinate

$$\tau = \lambda x^0 \tag{3, 10}$$

and assume that every field quantity is a function of (τ, x^1, x^2, x^3) rather than of (x^0, x^1, x^2, x^3). If φ is such a quantity we now assume that φ, $\varphi_{|m}$ and $\partial\varphi/\partial\tau$ are of the same order of magnitude, so that $\varphi_{|0}$ is of the order of $\lambda\varphi$.

From this we conclude that if T_{00} in (3, 7) is of the order of magnitude of λ^q, then T_{0n} will be of the order of λ^{q+1} and T_{mn} of the order of λ^{q+2}.

Further, it follows from well-known considerations concerning the first approximation (the conservation of energy for the motion of a point) that γ_{00}, which is the potential energy of a unit mass, is of the same order of magnitude as the square of the velocity and is thus, in our present notation, of the order of λ^2. Hence we have the following orders of magnitude for the γ's:

$$\gamma_{00} \sim \lambda^2; \qquad \gamma_{0n} \sim \lambda^3; \qquad \gamma_{mn} \sim \lambda^4. \tag{3, 11}$$

If we expand the γ's as power series in λ we must therefore take the lowest powers of the expansions to be of the orders given in (3, 11). The fact that only second derivatives of the γ's with respect to the time enter the equation (3, 5) shows that the powers of λ in successive terms of the expansions of the γ's may differ by two.[6] We are thus led to the simple assumption that

$$\begin{aligned} \gamma_{00} &= \lambda^2 \underset{2}{\gamma_{00}} + \lambda^4 \underset{4}{\gamma_{00}} + \lambda^6 \underset{6}{\gamma_{00}} + \cdots, \\ \gamma_{0n} &= \lambda^3 \underset{3}{\gamma_{0n}} + \lambda^5 \underset{5}{\gamma_{0n}} + \cdots, \\ \gamma_{mn} &= \lambda^4 \underset{4}{\gamma_{mn}} + \lambda^6 \underset{6}{\gamma_{mn}} + \cdots . \end{aligned} \tag{3, 12}$$

[6] The omission of terms with λ^{2l+1} in γ_{00}, γ_{mn} and with λ^{2l} in γ_{0m} is possible and natural, but logically not strictly necessary. The addition of the omitted terms of (3, 12) could be made in such a way that it would correspond to an introduction of a retarded potential (outgoing wave). Such a procedure would however, be artificial though it would not influence the equations of motion derived in II, as will be shown elsewhere.

We cannot discuss the question of convergence in general, but it is of interest to show that the new method of approximation can give convergent results even where this would not at first be expected. We consider the case of the one-dimensional wave equation in its simplest form

$$f_{xx} - f_{tt} = 0. \tag{3, 13}$$

If, in accordance with the main idea of the new method of approximation, we write

$$\begin{aligned} f &= \underset{0}{f} + \lambda^2 \underset{2}{f} + \lambda^4 \underset{4}{f} + \cdots, \\ f_{xx} &= \underset{0}{f_{xx}} + \lambda^2 \underset{2}{f_{xx}} + \lambda^4 \underset{4}{f_{xx}} + \cdots, \\ f_{tt} &= \lambda^2 f_{\tau\tau} = \lambda^2 \underset{0}{f_{\tau\tau}} + \lambda^4 \underset{2}{f_{\tau\tau}} + \lambda^6 \underset{4}{f_{\tau\tau}} + \cdots, \end{aligned} \tag{3, 14}$$

we obtain from (3, 13) the successive equations

$$\underset{0}{f_{xx}} = 0, \tag{3, 15a}$$

$$\underset{2}{f_{xx}} - \underset{0}{f_{\tau\tau}} = 0, \tag{3, 15b}$$

$$\underset{4}{f_{xx}} - \underset{2}{f_{\tau\tau}} = 0, \tag{3, 15c}$$

From these equations we can find the general solution of the wave equation (3, 13) expressed as a power series in λ. For simplicity we shall consider only the case of a sinusoidal wave so that, out of the totality of solutions of (3, 15a),

$$\underset{0}{f} = A(\tau) + xB(\tau), \tag{3, 16}$$

we choose the particular solution[7]

$$\underset{0}{f} = \sin \tau \tag{3, 17a}$$

and at each subsequent stage of the procedure we ignore all arbitrary functions which may enter. From (3, 15b), (3, 15c), . . . , we thus find

$$\underset{2}{f} = -\frac{x^2}{2!} \sin \tau, \tag{3, 17b}$$

$$\underset{4}{f} = \frac{x^4}{4!} \sin \tau, \tag{3, 17c}$$

so that the solution takes the form

$$f = \sin \tau \left\{ 1 - \frac{(x\lambda)^2}{2!} + \frac{(x\lambda)^4}{4!} - \cdots \right\} = \cos (\lambda x) \sin \tau.$$

---

[7] The inclusion of the solution $\underset{0}{f} = x \sin \tau$ also leads to sinusoidal waves, as is easily seen.

On replacing τ by λt we have

$$f = \cos(\lambda x)\sin(\lambda t) \tag{3, 18}$$

which is an exact solution of (3, 13).

4. **Expansion Properties of Field Quantities.** We shall show in this section that there is a simple general rule concerning the types of expansion which will occur when we treat the gravitational equations by the present method of approximation. This rule is that

Any component having an odd number of zero suffixes will have only odd powers of λ in its expansion, while any component having an even number of such suffixes will involve only even powers of λ in its expansion.

The fundamental equations (3, 12) show that the $\gamma_{\mu\nu}$ conform to this rule. The relations (1, 19), (1, 20), (1, 21) between $\gamma_{\mu\nu}$ and $h_{\mu\nu}$ have inverse relations of precisely the same form with γ and h interchanged, as

$$h_{00} = \tfrac{1}{2}\gamma_{00} + \tfrac{1}{2}\gamma_{ll}\,, \tag{4, 1}$$

$$h_{0n} = \gamma_{0n}\,, \tag{4, 2}$$

$$h_{mn} = \gamma_{mn} - \tfrac{1}{2}\delta_{mn}\gamma_{ll} + \tfrac{1}{2}\delta_{mn}\gamma_{00}\,, \tag{4, 3}$$

and from (3, 12) it follows that the expansions for the h's in powers of λ are of the form

$$\begin{aligned} h_{00} &= \lambda^2 \underset{2}{h}_{00} + \lambda^4 \underset{4}{h}_{00} + \lambda^6 \underset{6}{h}_{00} + \cdots, \\ h_{0n} &= \lambda^3 \underset{3}{h}_{0n} + \lambda^5 \underset{5}{h}_{0n} + \cdots, \\ h_{mn} &= \lambda^2 \underset{2}{h}_{mn} + \lambda^4 \underset{4}{h}_{mn} + \lambda^6 \underset{6}{h}_{mn} + \cdots, \end{aligned} \tag{4, 4}$$

showing that the h's also conform to the general rule.

Further, since the $\eta_{\mu\nu}$ trivially conform because η_{0n} vanishes, it follows from (1, 4) that the $g_{\mu\nu}$ also conform.

We may write the relation

$$g_{\mu\nu}g^{\nu\sigma} = \delta_\mu^\sigma \tag{4, 5}$$

in the form

$$g_{\mu n}g^{n\sigma} + g_{\mu 0}g^{0\sigma} = \delta_\mu^\sigma. \tag{4, 6}$$

The two groups of terms on the left differ by an even number of zero suffixes so that, since the δ_μ^σ trivially conform to the general rule, we shall obtain enough equations at each approximation for finding the expansions of the $g^{\mu\nu}$ if we assume that the general rule is valid for these components too. However, the $g^{\mu\nu}$ are uniquely determined in terms of the $g_{\mu\nu}$ by (4, 5) so that the expansions according to the general rule will give the only solution and extraneous powers of λ will necessarily have zero coefficients. Thus the rule is applicable to the $g^{\mu\nu}$ and so, also, to the $h^{\mu\nu}$.

Let us consider next the Christoffel symbols of both kinds. We have

$$[\mu\nu, \sigma] = \tfrac{1}{2}(g_{\nu\sigma|\mu} + g_{\sigma\mu|\nu} - g_{\mu\nu|\sigma}), \tag{4, 7}$$

and since the operation "$_{|0}$" introduces a factor λ while the operations "$_{|m}$" leave the order of magnitude unchanged it is evident that the fact that the $g_{\mu\nu}$ obey the general rule implies that the $[\mu\nu, \sigma]$ do too.

The Christoffel symbols of the second kind are defined by

$$\begin{Bmatrix} \lambda \\ \mu\nu \end{Bmatrix} = g^{\lambda\sigma}[\mu\nu, \sigma] \tag{4, 8}$$

and since whenever we have a dummy suffix we shall have either no extra zero suffixes or two such suffixes entering any term in the implied summation, the fact that $g^{\lambda\sigma}$ and $[\mu\nu, \sigma]$ separately conform to the general rule shows that this is true also of the $\begin{Bmatrix} \lambda \\ \mu\nu \end{Bmatrix}$.

In the course of the above considerations we have shown that neither the entry of dummy suffixes nor the operations "$_{|m}$", "$_{|0}$" disturbs the operation of the general rule. It follows that if, by the use of these operations alone, we form new quantities from quantities which conform to the rule these new quantities must also obey the general rule. This has already been exemplified by our discussion of the Christoffel symbols, and since all the quantities we shall have to consider, such as

$$R(= g^{\mu\nu} R_{\mu\nu}), \quad \Lambda_{\mu\nu}, \text{ etc.}$$

are new quantities of this type, we see that all the quantities with which we have to deal will have expansions in powers of λ whose general character is summed up in the statement at the head of this section.

5. **Alternative Form of the Equations When Singularities Are Absent.** In this section and the next we shall discuss the case where no singularities are present in the field. This case is, of course, trivial from the physical point of view since it corresponds to the complete absence of matter and, indeed, according to our method of approximation leads to the Galilean solution. Despite this, the discussion of this case will not be without value, for it will serve to exhibit the general mechanism of the theory and will form a convenient introduction to the later, more difficult discussion necessary when singularities are present.

Let us summarize some of the results we have obtained so far. The field has been subjected to the two restrictions

I The Gravitational Field Equations, and

II The Coördinate Conditions,

from which we have found

III The Surface Integral Conditions.

That is, if we take coördinate conditions

$$\gamma_{00|0} - \gamma_{0n|n} = 0, \tag{1, 25}$$

$$\gamma_{nm|n} = 0, \tag{1, 26}$$

the field equations take the form

$$\gamma_{00|ss} = 2\Lambda_{00}, \tag{1, 30}$$

$$\gamma_{0n|ss} = 2\Lambda_{0n}, \tag{1, 31}$$

$$\gamma_{mn|ss} = 2\Lambda_{mn}, \tag{1, 32}$$

and from these two groups of equations we obtain the surface integral conditions

$$\int (\gamma_{00|0n} - 2\Lambda_{0n}) \cos (\mathbf{n}\cdot\mathbf{N})\, dS = 0, \tag{2, 7}$$

$$\int 2\Lambda_{mn} \cos (\mathbf{n}\cdot\mathbf{N})\, dS = 0, \tag{2, 8}$$

and also the results

$$(\gamma_{00|0n} - 2\Lambda_{0n})_{|n} = 2\Lambda_{00|0} - 2\Lambda_{0n|n} = 0, \tag{2, 9}$$

$$2\Lambda_{mn|n} = 0, \tag{2, 10}$$

which are essential for the validity of the surface integral conditions for arbitrary surfaces.

We shall now show that the following two sets of equations (5, 1), (5, 2) are equivalent when no singularities are present.

(5, 1)		(5, 2)	
(a)	$\gamma_{00\|ss} = 2\Lambda_{00}$,	(a)	$\gamma_{00\|ss} = 2\Lambda_{00}$,
(b)	$\gamma_{0n\|ss} = 2\Lambda_{0n}$,	(b)	$\gamma_{0n\|ss} = 2\Lambda_{0n}$,
(c)	$\gamma_{00\|0} - \gamma_{0n\|n} = 0$;	(c′)	$\Lambda_{00\|0} - \Lambda_{0n\|n} = 0$,
		(c″)	$\int (\gamma_{00\|0n} - 2\Lambda_{0n}) \cos (\mathbf{n}\cdot\mathbf{N})\, dS = 0$;
(d)	$\gamma_{mn\|ss} = 2\Lambda_{mn}$,	(d)	$\gamma_{mn\|ss} = 2\Lambda_{mn}$,
(e)	$\gamma_{mn\|n} = 0$,	(e′)	$\Lambda_{mn\|n} = 0$,
		(e″)	$\int 2\Lambda_{mn} \cos (\mathbf{n}\cdot\mathbf{N})\, dS = 0$.

In (5, 1) we have merely the field equations and coördinate conditions and we show essentially that the coördinate conditions may be replaced by the surface integral conditions[8] and the conditions (2, 9) (2, 10). The proof for the present

[8] When singularities are absent (5, 2c′), (5, 2c″) and also (5, 2e′), (5, 2e″) are equivalent equations, but we include them all here in order to facilitate comparison with the situation which arises when singularities are present.

case is trivial. For we have already shown that (5, 1) implies (5, 2) and the converse follows at once from the following considerations.

From (5, 2a) (5, 2b) and (5, 2c′) we have

$$(\gamma_{00|0} - \gamma_{0n|n})_{|ss} = 2\Lambda_{00|0} - 2\Lambda_{0n|n} = 0,$$

and since there are no singularities and the γ's must be zero at infinity this gives

$$\gamma_{00|0} - \gamma_{0n|n} = 0$$

which is (5, 1c). The proof for γ_{mn} is similar.

6. **Splitting of the Equations When Singularities Are Absent.** In the first section we gave a prescription for separating the terms of each of the field equations into two well-defined groups. In this section we shall discuss the splitting of the gravitational equations according to powers of λ and shall show why just this method of separation is implied by our method of approximation.

It is necessary first to introduce certain notations. Consider the quantity

$$h_{mn|0s}\,. \tag{6, 1}$$

When h_{mn} is expanded in powers of λ we write

$$h_{mn} = \lambda^2 \underset{2}{h}_{mn} + \lambda^4 \underset{4}{h}_{mn} + \cdots + \lambda^{2l} \underset{2l}{h}_{mn} + \cdots, \tag{6, 2}$$

where the numbers underneath the h's on the right serve the double purpose of distinguishing between the different functions h on the right and of showing with what power of λ each is associated in the expansion.

Now the fundamental assumption of our method of approximation requires that h_{mn} be a function of $(\lambda x^0, x^1, x^2, x^3)$ so that

$$h_{mn|s} = \frac{\partial h_{mn}}{\partial x^s}$$

but

$$h_{mn|0} = \frac{\partial h_{mn}}{\partial x^0} = \lambda\,\frac{\partial h_{mn}}{\partial \tau}.$$

In order to distinguish between ordinary differentiation with respect to (x^0, x^1, x^2, x^3) and ordinary differentiation with respect to (τ, x^1, x^2, x^3) we shall denote the latter by a comma followed by an appropriate suffix:

$$h_{mn|s} = \frac{\partial h_{mn}}{\partial x^s} = h_{mn,s}\,, \tag{6, 3}$$

$$h_{mn|0} = \frac{\partial h_{mn}}{\partial x^0} = \lambda\,\frac{\partial h_{mn}}{\partial \tau} = \lambda\, h_{mn,0}\,. \tag{6, 4}$$

Thus h_{mn}, $h_{mn,s}$ and $h_{mn,0}$ are all of the same order, but $h_{mn|0}$ belongs to a power of λ one higher.

With this convention we may write the expansion of (6, 1) in the form

$$(6,\ 5)\qquad h_{mn|0s} = \lambda h_{mn,0s} = \lambda^3 \underset{2}{h}_{mn,0s} + \lambda^5 \underset{4}{h}_{mn,0s} + \cdots + \lambda^{2l+1} \underset{2l}{h}_{mn,0s} + \cdots.$$

Now, however, the number underneath each h on the right no longer indicates directly the power of λ with which it is associated. We therefore write a 1 underneath each zero suffix following a comma for every h having a number underneath so that (6, 5) becomes

$$(6,\ 6)\qquad h_{mn|0s} = \lambda h_{mn,0s} = \lambda^3 \underset{2}{h}_{mn,\underset{1}{0}s} + \lambda^5 \underset{4}{h}_{mn,\underset{1}{0}s} + \cdots + \lambda^{2l+1} \underset{2l}{h}_{mn,\underset{1}{0}s} + \cdots.$$

Thus now the sum of the numbers underneath each h gives the power of λ with which it is associated while the first of these numbers indicates the particular function h we are considering. This notation is then consistent with the natural notation for a product of h's.

We consider now what happens when we introduce the power series expansions for the h's in the equations (1, 27), (1, 28), (1, 29). On equating to zero the coefficients of the various powers of λ we shall obtain

$$(6,\ 7)\qquad \underset{2l}{h}_{00,ss} = 2\underset{2l}{L}_{00},$$

$$(6,\ 8)\qquad \underset{2l+1}{h}_{0n,ss} = 2\underset{2l+1}{L}_{0n},$$

$$(6,\ 9)\qquad \underset{2l}{h}_{mn,ss} = 2\underset{2l}{L}_{mn}.$$

The lowest h's are $\underset{2}{h}_{00}$, $\underset{3}{h}_{0n}$, and $\underset{2}{h}_{mn}$, and these will therefore be the quantities determined in the first approximation. They correspond to $l = 1$ in the scheme of (6, 7) (6, 8) (6, 9). Thus at any stage, say l, the quantities to be determined are $\underset{2l}{h}_{00}$, $\underset{2l+1}{h}_{0n}$, $\underset{2l}{h}_{mn}$, and the quantities already known from the solutions of the previous approximations are the h's having lower numbers underneath.

But if we look at the forms of the L's, as given in (1, 15), (1, 16), (1, 17) we see that at the stage l we have either quadratic terms or linear terms involving differentiations with respect to x^0. The quadratic terms can only involve h's of lower order than for l, and the linear terms may be written as

$$(6,\ 10)\qquad \underset{2l-1}{h}_{0s,\underset{1}{0}s} - \tfrac{1}{2}\underset{2l-2}{h}_{ss,\underset{1}{0}\underset{1}{0}} \quad \text{in } \underset{2l}{L}_{00},$$

$$(6,\ 11)\qquad \text{none} \quad \text{in } \underset{2l+1}{L}_{0n},$$

$$(6,\ 12)\qquad \tfrac{1}{2}\underset{2l-2}{h}_{mn,\underset{1}{0}\underset{1}{0}} - \tfrac{1}{2}\underset{2l-1}{h}_{0m,\underset{1}{0}n} - \tfrac{1}{2}\underset{2l-1}{h}_{0n,\underset{1}{0}m} \quad \text{in } \underset{2l}{L}_{mn}.$$

These are all known functions from the previous approximations. Thus the whole of $\underset{2l}{L_{00}}$, $\underset{2l+1}{L_{0n}}$, $\underset{2l}{L_{mn}}$ for given l are known from the solutions of the previous approximations. This is the reason for the particular method of separation of the field equations into two parts described in 1. When the separation is made in this manner and the power series expansions are inserted for the h's in (1, 27) (1, 28), (1, 29), for each power of λ the corresponding coefficients automatically group themselves into those quantities which enter for the first time with the approximation in question and those which are already known, at least in principle, from the previous approximations. These two groups correspond exactly to the left and right hand sides of (1, 27), (1, 28), (1, 29).

Before we can solve the approximation equations we must also split the coördinate conditions (1, 25), (1, 26), and the relations between the h's and γ's according to powers of λ. It turns out that we may take at each stage

$$\underset{2l}{\gamma_{00,ss}} = 2\underset{2l}{\Lambda_{00}}, \qquad \underset{2l+1}{\gamma_{0n,ss}} = 2\underset{2l+1}{\Lambda_{0n}}, \qquad \underset{2l\ \ 1}{\gamma_{00,0}} - \underset{2l+1}{\gamma_{0n,n}} = 0; \tag{6, 13}$$

$$\underset{2l}{\gamma_{mn,ss}} = 2\underset{2l}{\Lambda_{mn}}, \qquad \underset{2l}{\gamma_{mn,n}} = 0, \tag{6, 14}$$

where the Λ's are known because of the solutions of the previous approximations.

We may also split the alternative equations (5, 2) and use, instead, at each stage

$$\begin{gathered} \underset{2l}{\gamma_{00,ss}} = 2\underset{2l}{\Lambda_{00}}, \qquad \underset{2l+1}{\gamma_{0n,ss}} = 2\underset{2l+1}{\Lambda_{0n}}, \qquad \underset{2l\ \ 1}{\Lambda_{00,0}} - \underset{2l+1}{\Lambda_{0n,n}} = 0, \\ \int \left(\underset{2l\ \ 1}{\gamma_{00,0n}} - 2\underset{2l+1}{\Lambda_{0n}}\right) \cos\,(\mathbf{n}\cdot\mathbf{N})\,dS; \end{gathered} \tag{6, 15}$$

$$\underset{2l}{\gamma_{mn,ss}} = 2\underset{2l}{\Lambda_{mn}}, \qquad \underset{2l}{\Lambda_{mn,n}} = 0, \qquad \int 2\underset{2l}{\Lambda_{mn}} \cos\,(\mathbf{n}\cdot\mathbf{N})\,dS = 0. \tag{6, 16}$$

As in the case of the unsplit equations, the surface integral conditions are consequences of the others because of the absence of singularities, and the whole splitting actually presents no fundamental difficulties for this case.

7. The General Theory When Singularities Are Present. The existence of singularities in the field introduces certain factors which make the theory developed for the regular case inadequate. For, although the equations of the field are undefined at the singularities, their validity in the regular region is sufficient to determine the motion of these singularities. The slightest alteration in the position of a singularity amounts to an arbitrarily large alteration for a point near enough to the singularity, and we are therefore not permitted to make use of approximate expressions for the equations of motion in the development of our method of approximation. This fact leads to a new difficulty, in the approximation method, which must be discussed more fully.

Let there be p particles producing the field. We may represent their positions at any time by means of their spatial coördinates $\overset{k}{\xi}{}^m(\tau)$, $k = 1, 2, \cdots, p$. At these points the field will be singular, but we may enclose each of the singularities within a small surface,[9] and then the region exterior to these p surfaces will be regular.

Although the equations (5, 1), (5, 2) are undefined at the singularities, they have meaning in the regular region and we shall show that they can still be regarded as in some sense equivalent. The discussion can be divided into two parts, one dealing with the (a), (b) and (c) equations, which involve the suffix zero, and the other with the remaining equations having only spatial suffixes. We consider the latter. The essential structure of the (d) and (e) equations is preserved if we omit the suffix m and write for the total field

(7, 1)		(7, 2)	
(d)	$\gamma_{n\|ss} = 2\Lambda_n,$	(d)	$\gamma_{n\|ss} = 2\Lambda_n,$
(e)	$\gamma_{n\|n} = 0,$	(e′)	$\Lambda_{n\|n} = 0,$
		(e″)	$\int 2\Lambda_n \cos(\mathbf{n}\cdot\mathbf{N})\, dS = 0.$

The proof that (7, 1) implies (7, 2) has already been given in essence in 2. To prove the converse we first obtain from (7, 2d)

$$\gamma_{n|nss} = 2\Lambda_{n|n}, \tag{7, 3}$$

this being valid outside the surfaces enclosing the singularities. To solve this we make an analytic continuation of the functions Λ_n into the interiors of these surfaces in such a way that $\Lambda_{n|n}$ is everywhere zero. This is certainly possible because of the validity of (7, 2e″). So (7, 3) now becomes

$$\gamma_{n|nss} = 0$$

which, being everywhere valid, has the unique solution

$$\gamma_{n|n} = 0$$

which is (7, 1e).

Thus we have shown that if we make an analytic continuation of Λ_n so that (7, 2e′) is valid everywhere, then (7, 1) and (7, 2) are equivalent outside the surfaces enclosing the singularities.

It is clear from the proof that the result will hold for any surfaces enclosing the singularities.

For the (a), (b) and (c) equations a similar proof can also be given. In this case it is necessary to make an analytic continuation of the quantities Λ_{00} and

[9] Throughout the argument we assume that we are dealing with the situation at some definite time τ, allowing time to flow again only after the argument is concluded.

Λ_{0n} in such a way that (5, 2c′) is valid everywhere, this being possible because of (5, 2c″). We omit the details of this part of the proof that (5, 1), (5, 2) may be considered as equivalent even where singularities are present and shall regard the proof as complete.

To show the difficulty brought in by the use of our approximation method let us now consider only the equations (7, 2d), (7, 2e′) omitting the surface integral (7, 2e″). These equations determine the field in each of the approximation steps if the motions of the singularities are prescribed. The motion of the particles is then arbitrary as, for example, in the electrodynamical problem and the field is determined in each of the approximation steps by the equations

$$\underset{2l}{\gamma_{n,ss}} = 2\underset{2l}{\Lambda_n}$$

$$\underset{2l}{\Lambda_{n,n}} = 0.$$

The contradiction is evident if we try to add to these equations the surface condition split according to our approximation method. We then have the additional equation

$$\int^{k} 2\underset{2l}{\Lambda_n} \cos (\mathbf{n}\cdot\mathbf{N}) = 0 \tag{7, 4}$$

where (k) on top of an integral sign means that the surface of integration encloses only the k-singularity. We have in (7, 4) an infinite set of equations containing the functions $\overset{k}{\xi}$ and their time derivatives. These equations cannot be satisfied by the arbitrarily given $\overset{k}{\xi}$ functions characterising the motion.

This also shows how the difficulty can be avoided. We have to consider instead of (5, 1) or (5, 2) a more general set of conditions governing the field which contains those equations as a particular case. Since it is the surface integral conditions which cause the trouble we remove (5, 2c″), (5, 2e″) from the set (5, 2) and consider the significance of what remains.

In making this generalisation we have, of course, gone beyond the gravitational equations to others which contain them as a special case, and we must now discuss what changes have been induced in (5, 1) by this generalisation.

Since the surface integrals are independent of the surfaces, their values will be functions of the time alone through the ξ's and their derivatives. There is therefore no loss of generality if we denote these integrals taken over the p surfaces enclosing the various singularities by $4\pi\overset{k}{c}_0(\tau)$, $4\pi\overset{k}{c}_m(\tau)$:

$$\begin{aligned} \frac{1}{4\pi}\int^{k} (\gamma_{00|0} - 2\Lambda_{0n}) \cos (\mathbf{n}\cdot\mathbf{N})\, dS &= \overset{k}{c}_0(\tau), \\ \frac{1}{4\pi}\int^{k} 2\Lambda_{mn} \cos (\mathbf{n}\cdot\mathbf{N})\, dS &= \overset{k}{c}_m(\tau). \end{aligned} \tag{7, 5}$$

With this notation we shall now prove that the following two sets of equations (7, 6), (7, 7) are equivalent in a certain sense which will be explained in the course of the proof:

	(7, 6)		(7, 7)
(a)	$\gamma_{00\|ss} = 2\Lambda_{00}$,	(a)	$\gamma_{00\|ss} = 2\Lambda_{00}$,
(b)	$\gamma_{0n\|ss} = 2\Lambda_{0n}$,	(b)	$\gamma_{0n\|ss} = 2\Lambda_{0n}$,
(c)	$\gamma_{00\|0} - \gamma_{0n\|n} = -\sum_{k=1}^{p} \left\{ \overset{k}{c}_0 / \overset{k}{r} \right\}$;	(c′)	$\Lambda_{00\|0} - \Lambda_{0n\|n} = 0$;
(d)	$\gamma_{mn\|ss} = 2\Lambda_{mn}$,	(d)	$\gamma_{mn\|ss} = 2\Lambda_{mn}$,
(e)	$\gamma_{mn\|n} = -\sum_{k=1}^{p} \left\{ \overset{k}{c}_m / \overset{k}{r} \right\}$,	(e′)	$\Lambda_{mn\|n} = 0$.

Here $\overset{k}{r}$ is the "distance" from x^n to the k-singularity:

$$\overset{k}{r} = \left[\left(x^s - \overset{k}{\xi}{}^s \right) \left(x^s - \overset{k}{\xi}{}^s \right) \right]^{\frac{1}{2}}. \tag{7, 8}$$

We may introduce the surfaces enclosing the singularities as before and these equations will certainly have meaning outside them. The proof of their equivalence can here too be broken up into two parts and we shall only prove the equivalence for the (d) and (e) parts. Omitting the suffix m as before, we have

	(7, 9)		(7, 10)
(d)	$\gamma_{n\|ss} = 2\Lambda_n$,	(d)	$\gamma_{n\|ss} = 2\Lambda_n$,
(e)	$\gamma_{n\|n} = -\sum_{k=1}^{p} \left\{ \overset{k}{c} / \overset{k}{r} \right\}$,	(e′)	$\Lambda_{n\|n} = 0$,

with the notation

$$\frac{1}{4\pi} \int^{k} 2\Lambda_n \cos (\mathbf{n} \cdot \mathbf{N}) \, dS = \overset{k}{c}(\tau). \tag{7, 11}$$

We begin by proving that (7, 10) implies (7, 9) under certain conditions of continuation. It is no longer possible to make an analytic continuation of Λ_n in such a way that (e′) is everywhere satisfied since this would imply that the surface integral is necessarily zero. In fact, from (7, 11) we see, by Gauss' theorem, that the continuation must be such that

$$\frac{1}{4\pi} \int^{k} 2\Lambda_{n|n} \, dv = \frac{1}{4\pi} \int^{k} 2\Lambda_n \cos (\mathbf{n} \cdot \mathbf{N}) \, dS = \overset{k}{c}(\tau). \tag{7, 12}$$

It is simplest for our purposes to make the continuation in such a way that Λ_n and $\Lambda_{n|n}$ are continuous at the surfaces, and that $\Lambda_{n|n}$ has a constant sign inside each surface and satisfies (7, 12).

Such a continuation is possible for any surfaces surrounding the singularities and can be made in such a way that when these surfaces shrink to zero size the function $\Lambda_{n|n}$ goes over to a sum of Dirac δ-functions:

$$\Lambda_{n|n} \rightarrow 4\pi \sum_{k=1}^{p} \overset{k}{c}(\tau)\cdot\delta\left(x^1 - \overset{k}{\xi}{}^1\right)\delta\left(x^2 - \overset{k}{\xi}{}^2\right)\delta\left(x^3 - \overset{k}{\xi}{}^3\right). \tag{7, 13}$$

From (7, 10d) we have now

$$\gamma_{n|nss} = 2\Lambda_{n|n}$$

so that

$$\gamma_{n|n}(x) = -\frac{1}{4\pi}\int \frac{2\Lambda_{n|n}(x')}{r(x, x')}\, dv', \tag{7, 14}$$

where the integral is to be taken over the whole domain of x^n and $r(x, x')$ is the "distance" from x^n to x'^n:

$$r(x, x') = [(x^s - x'^s)(x^s - x'^s)]^{\frac{1}{2}}. \tag{7, 15}$$

Because of the validity of (7, 10e′) outside the surfaces we may write (7, 14) as

$$\gamma_{n|n}(x) = -\frac{1}{4\pi}\sum_{k=1}^{p}\int^{k} \frac{2\Lambda_{n|n}(x')}{r(x, x')}\, dv', \tag{7, 16}$$

the integrals being taken only over the interiors of the surfaces. On shrinking these surfaces we may regard $r(x, x')$ as constants over the various domains of integration and write

$$\gamma_{n|n}(x) = -\frac{1}{4\pi}\sum_{k=1}^{p} 1/\left(\overset{k}{r}(x)\right)\int^{k} 2\Lambda_{n|n}\, dv,$$

and by (7, 12) this is

$$\gamma_{n|n} = -\sum_{k=1}^{p}\left\{\overset{k}{c}(\tau)/\overset{k}{r}\right\}$$

which is (7, 9e).

We have therefore shown that with the analytic continuation used above the equations (7, 10) imply the equations (7, 9).

To prove the converse we form from (7, 9) the relation

$$(\gamma_{n|s} - \gamma_{n|n})_{|s} = 2\Lambda_n + \sum_{k=1}^{p}\left\{\overset{k}{c}(\tau)/\overset{k}{r}\right\}_{|n}. \tag{7, 17}$$

If we now form the surface integrals of the "normal" components of the two sides of this equation for each of the surfaces enclosing the singularities in turn, the left hand side will give zero, as explained in 2, and we shall have left

$$\begin{aligned}\int^{k} 2\Lambda_n \cos(\mathbf{n}\cdot\mathbf{N})\, dS &= -\int^{k}\left\{\overset{k}{c}/\overset{k}{r}\right\}_{|n}\cos(\mathbf{n}\cdot\mathbf{N})\, dS \\ &= -\overset{k}{c}\int^{k}\left\{1/\overset{k}{r}\right\}_{|n}\cos(\mathbf{n}\cdot\mathbf{N})\, dS = 4\pi\overset{k}{c}(\tau)\end{aligned}$$

which is (7.11).

That the validity of (7, 10e′) for the regular region is contained in (7, 9) is trivial, and the equivalence of (7, 6d, e) and (7, 7d, e′) is therefore proved.

The proof of the equivalence for the remaining equations of (7, 6), (7, 7) containing the suffix zero presents no essentially new problems and will be omitted.

The whole point of our elaborate procedure in writing all the equations of the field in two equivalent forms is now clear since the present generalisation from (5, 1) to (7, 6) could not be made in a convincing manner without the aid of the parallelism with (5, 2) and (7, 7).

Owing to the absence of the surface integral conditions in (7, 7), there is no longer any objection to the application of our method of approximation to the solution of this set of equations. The λ's will cause a splitting of the equations just as before, except that the surface integral conditions will be absent. However, at each stage we may write

(7, 18)
$$\frac{1}{4\pi}\int^{k}\Big(\underset{2l\ \ \ 1}{\gamma_{00,0n}} - 2\underset{2l+1}{\Lambda_{0n}}\Big)\cos(\mathbf{n}\cdot\mathbf{N})\,dS = \underset{2l+1}{\overset{k}{c_0}}(\tau),$$
$$\frac{1}{4\pi}\int^{k} 2\underset{2l}{\Lambda_{mn}}\cos(\mathbf{n}\cdot\mathbf{N})\,dS = \underset{2l}{\overset{k}{c_m}}(\tau),$$

and with this notation we have the result, in precisely the same manner as for (7, 6), (7, 7), that for each stage of the approximation the following sets of equations (7, 19), (7, 20) are equivalent:

(7, 19)		(7, 20)	
(a)	$\underset{2l}{\gamma_{00,ss}} = 2\underset{2l}{\Lambda_{00}}$,	(a)	$\underset{2l}{\gamma_{00,ss}} = 2\underset{2l}{\Lambda_{00}}$,
(b)	$\underset{2l+1}{\gamma_{0n,ss}} = 2\underset{2l+1}{\Lambda_{0n}}$,	(b)	$\underset{2l+1}{\gamma_{0n,ss}} = 2\underset{2l+1}{\Lambda_{0n}}$,
(c)	$\underset{2l\ \ 1}{\gamma_{00,0}} - \underset{2l+1}{\gamma_{0n,n}} = -\sum_{k=1}^{p}\Big\{\underset{2l+1}{\overset{k}{c_0}}(\tau)/\overset{k}{r}\Big\}$;	(c′)	$\underset{2l\ \ 1}{\Lambda_{00,0}} - \underset{2l+1}{\Lambda_{0n,n}} = 0$;
(d)	$\underset{2l}{\gamma_{mn,ss}} = 2\underset{2l}{\Lambda_{mn}}$,	(d)	$\underset{2l}{\gamma_{mn,ss}} = 2\underset{2l}{\Lambda_{mn}}$,
(e)	$\underset{2l}{\gamma_{mn,n}} = -\sum_{k=1}^{p}\Big\{\underset{2l}{\overset{k}{c_m}}(\tau)/\overset{k}{r}\Big\}$,	(e′)	$\underset{2l}{\Lambda_{mn,n}} = 0$.

In the actual solving of the equations it is simpler to work with the sets (7, 19) rather than with (7, 20). At each stage we have to solve equations of the type $\gamma_{,ss} = 2\Lambda$ and in order to make the whole solution unambiguous we must impose the conditions that the field shall be Galilean at infinity and that no harmonic functions of higher type than simple poles may be added to the partial solutions except insofar as their addition is forced by the coördinate conditions (7, 19c), (7, 19e). Let us suppose we have been able to solve all the successive approximations. Then the quantities $\overset{k}{c}_0(\tau)$, $\overset{k}{c}_m(\tau)$ are given by the relations

$$\overset{k}{c}_0(\tau) = \sum_{l=1}^{\infty} \lambda^{2l+1} \underset{2l+1}{\overset{k}{c}_0}(\tau), \tag{7, 21}$$

$$\overset{k}{c}_m(\tau) = \sum_{l=1}^{\infty} \lambda^{2l} \underset{2l}{\overset{k}{c}_m}(\tau). \tag{7, 22}$$

Our solution will not in general be a solution of the gravitational equations since (7, 6), (7, 7) are more general than those equations. However, if we now put

$$\overset{k}{c}_0(\tau) = 0, \qquad \overset{k}{c}_m(\tau) = 0 \tag{7, 23}$$

we impose such conditions on the motions of the singularities that our solutions will indeed become the solutions of the gravitational equations we are actually interested in.

The differential equations (7, 23) for the ξ's are really independent of λ since they must be expressed in terms, not of the auxiliary time τ but of the true time x^0, and when this is done the λ's will be necessarily reabsorbed.

In practice, of course, it is impossible to carry the computation beyond the first few stages. Let us suppose, then, that we have been able to solve the successive approximations up to some stage $l = q$. In this case, if we put

$$\sum_{l=1}^{q} \lambda^{2l+1} \underset{2l+1}{\overset{k}{c}_0}(\tau) = 0, \qquad \sum_{l=1}^{q} \lambda^{2l} \underset{2l}{\overset{k}{c}_m}(\tau) = 0, \tag{7, 24}$$

we shall obtain solutions of the gravitational equations correct to terms of the order $(2q + 1)$, and the equations (7, 24) will give the approximate equations of motion up to this order.

8. **The Zero Coördinate Condition.** We show in this section that the solution of our equations can always be made in such a way that

$$\underset{2l+1}{\overset{k}{c}_0}(\tau) = 0, \tag{8, 1}$$

thus showing that the conditions

$$\overset{k}{c}_0(\tau) = \sum_{l=1}^{\infty} \lambda^{2l+1} \underset{2l+1}{\overset{k}{c}_0}(\tau) \tag{7, 21}$$

place no restriction on the motion of the singularities. This result is of significance because the conditions (7, 22) are alone sufficient to describe the motion completely and any further condition, if not redundant, would cause an overdetermination of the motion.

We actually use (8, 1) as normalisation conditions for each stage of the approximation and they are essential for the uniqueness of the solution.

The significant equations for the present argument are

$$\underset{2l}{\gamma_{00,ss}} = 2\underset{2l}{\Lambda_{00}}, \tag{7, 19a}$$

$$\underset{2l+1}{\gamma_{0n,ss}} = 2\underset{2l+1}{\Lambda_{0n}}, \tag{7, 19b}$$

$$\underset{2l\ \ 1}{\gamma_{00,0}} - \underset{2l+1}{\gamma_{0n,n}} = -\sum_{k=1}^{p}\left\{\underset{2l+1}{\overset{k}{c_0}(\tau)}/\overset{k}{r}\right\}, \tag{7, 19c}$$

where the Λ's are known from the solutions of the previous approximations, and we shall suppose that we have a solution of these equations. If we introduce the quantities $\overset{k}{\Gamma}(\tau)$ by means of the equation

$$\underset{2l\ 1}{\overset{k}{\Gamma}_{,0}}(\tau) = \underset{2l+1}{\overset{k}{c_0}}(\tau), \tag{8, 2}$$

we may write (7, 19c) in the form

$$\begin{aligned}\underset{2l\ \ 1}{\gamma_{00,0}} - \underset{2l+1}{\gamma_{0n,n}} &= -\sum_{k=1}^{p}\left\{\left(\underset{2l}{\overset{k}{\Gamma}}/\overset{k}{r}\right)_{,\underset{1}{0}} - \underset{2l}{\overset{k}{\Gamma}}\left(1/\overset{k}{r}\right)_{,\underset{1}{0}}\right\} \\ &= -\sum_{k=1}^{p}\left\{\left(\underset{2l}{\overset{k}{\Gamma}}/\overset{k}{r}\right)_{,\underset{1}{0}} + \left(\underset{2l}{\overset{k}{\Gamma}}\ \underset{1}{\dot{\overset{k}{\xi}}{}^{n}}/\overset{k}{r}\right)_{,n}\right\},\end{aligned} \tag{8, 3}$$

where $\dot{\xi} = \dfrac{d\xi}{d\tau}$.

From (7, 19a), (7, 19b) we see that $\underset{2l}{\gamma_{00}}$ and $\underset{2l+1}{\gamma_{0n}}$ are arbitrary to the extent of additive harmonic functions and we may therefore add simple poles to them to form the new quantities

$$\underset{2l}{\gamma'_{00}} = \underset{2l}{\gamma_{00}} + \sum_{k=1}^{p}\left\{\underset{2l}{\overset{k}{\Gamma}}/\overset{k}{r}\right\}, \tag{8, 4}$$

$$\underset{2l+1}{\gamma'_{0n}} = \underset{2l+1}{\gamma_{0n}} - \sum_{k=1}^{p}\left\{\underset{2l}{\overset{k}{\Gamma}}\ \dot{\overset{k}{\xi}}{}^{n}/\overset{k}{r}\right\}. \tag{8, 5}$$

These new γ's however, while still satisfying (7, 19a), (7, 19b) will be such that

$$\underset{2l\ \ 1}{\gamma'_{00,0}} - \underset{2l+1}{\gamma'_{0n,n}} = 0. \tag{8, 6}$$

Since the c_0's now vanish, the surface integrals will also be zero and thus the zero coördinate condition will not affect the motion. This theorem and our previous results show us that the equations which must form the basis of the actual calculation of the field and the equations of motion of the singularities are:

$$\underset{2l}{\gamma_{00,ss}} = 2\underset{2l}{\Lambda_{00}}, \tag{8, 7a}$$

$$\underset{2l+1}{\gamma_{0n,ss}} = 2\underset{2l+1}{\Lambda_{0n}}, \tag{8, 7b}$$

$$\underset{2l}{\gamma_{00,\underset{1}{0}}} - \underset{2l+1}{\gamma_{0n,n}} = 0; \tag{8, 7c}$$

$$\underset{2l}{\gamma_{mn,ss}} = 2\underset{2l}{\Lambda_{mn}}, \tag{8, 7d}$$

and

$$\underset{2l}{\gamma_{mn,n}} = -\sum_{k=1}^{p} \left\{ \underset{2l}{\overset{k}{c}_m}(\tau)/\overset{k}{r} \right\}, \tag{8, 7e}$$

with

$$\underset{2l}{\overset{k}{c}_m}(\tau) = \frac{1}{4\pi} \int^{k} 2\underset{2l}{\Lambda_{mn}} \cos(\mathbf{n}\cdot\mathbf{N})\, dS. \tag{8, 8}$$

The approximate equations of motion for the stage $l = q$ are given by

$$\sum_{l=1}^{q} \lambda^{2l} \underset{2l}{\overset{k}{c}_m}(\tau) = 0. \tag{8, 9}$$

II. Application of the General Theory

Note. In the first part of this paper we developed the general theory of a new method for solving the equations of gravitation by successive approximation and for obtaining the equations of motion, in principle to any desired degree of accuracy. In the present part we deal with the actual application of this method, carrying the calculation to such a stage that the main deviation from the Newtonian laws of motion is determined.

Unfortunately, as the work proceeds, the calculations become more and more extensive involving a great amount of technical detail which can have no intrinsic interest. To give all these calculations explicitly here would be quite impracticable and we are obliged to confine ourselves to stressing the general ideas of the work and merely announcing the actual results. For the convenience of anyone who may be interested in the details of the calculation, however, the entire computation of this part of our paper has been deposited with the Institute for Advanced Study so as to be available for reference.[10]

9. **The Approximation $l = 1$.** The approximation $l = 0$ is trivial, leading to the Galilean case, and we proceed at once to the next approximation $l = 1$.

---

[10] c/o Secretary of the School of Mathematics, Institute for Advanced Study, Princeton, N. J. (U. S. A.).

Since the quantities $\underset{2}{\Lambda_{00}}$, $\underset{3}{\Lambda_{0n}}$, and, $\underset{2}{\Lambda_{mn}}$ are all zero and, as explained in 3, the $\underset{2}{\gamma_{mn}}$ are also zero, we have left from all the equations (8.7a), . . . , (8.7e), merely

$$\underset{2}{\gamma_{00,ss}} = 0, \tag{9.1}$$

$$\underset{3}{\gamma_{0n,ss}} = 0, \tag{9.2}$$

$$\underset{2}{\gamma_{00,0}} \underset{1}{} - \underset{3}{\gamma_{0n,n}} = 0. \tag{9.3}$$

The character of our whole solution will depend essentially upon the choice of the harmonic function we take as the solution of (9.1). We shall assume that the particles we are interested in have spherical symmetry and that the field is Galilean at infinity. In this case the solution of (9.1) is unique since each singularity in $\underset{2}{\gamma_{00}}$ must now, by (9.1) be a simple pole. We therefore have for $\underset{2}{\gamma_{00}}$ the solution

$$\underset{2}{\gamma_{00}} = 2\varphi, \qquad \varphi = \sum_{k=1}^{p} \left\{ -2\overset{k}{m}/\overset{k}{r} \right\}, \qquad \overset{k}{r} = \left[\left(x^s - \overset{k}{\xi^s} \right) \left(x^s - \overset{k}{\xi^s} \right) \right]^{\frac{1}{2}}, \tag{9.4}$$

where the p quantities $\overset{k}{m}$ are independent of the spatial coördinates x^s, and can depend at most only on the time.

From (9.2) we see that $\underset{3}{\gamma_{0n}}$ is also a harmonic function, and to determine it more exactly we must use the coördinate (9.3). From (9.3), (9.4) we have

$$\begin{aligned} \underset{3}{\gamma_{0n,n}} = \underset{2}{\gamma_{00,0}}\underset{1}{} &= \sum_{k=1}^{p} \left(-4\overset{k}{m}/\overset{k}{r} \right)_{,0} \\ &= \sum_{k=1}^{p} \left\{ \left(4\overset{k}{m}/\overset{k}{r} \right) \overset{k}{\dot{\xi}^n} \right\}_{,n} - \sum_{k=1}^{p} \left(-4\overset{k}{\dot{m}}/\overset{k}{r} \right). \end{aligned}$$

This equation can be solved without introducing new singularities only if $\overset{k}{\dot{m}} = 0$. In other words, the quantities $\overset{k}{m}$, which actually measure the masses of the point singularities, are necessarily constants. It is now evident that, under our general restricting conditions, $\underset{3}{\gamma_{0n}}$ is uniquely determined:

$$\underset{3}{\gamma_{0n}} = \sum_{k=1}^{p} \left\{ \left(4\overset{k}{m}/\overset{k}{r} \right) \overset{k}{\dot{\xi}^n} \right\}. \tag{9.5}$$

In all that follows we shall limit our considerations to the case of only two particles. This places no essential restriction on the results as far as the end of 15, their generalisation to p particles being trivial, and it permits a useful simplification of the rather inconvenient notation used for the general case.

For the case of two particles we shall write:

$$\text{(9.6)}\quad \begin{aligned} &\text{(a)} \quad -2\overset{1}{m}/\overset{1}{r} = \psi, \qquad -2\overset{2}{m}/\overset{2}{r} = \chi; \\ &\text{(b)} \quad \varphi = \psi + \chi; \\ &\text{(c)} \quad \overset{1}{\xi}{}^{s} = \eta^{s}, \qquad \overset{2}{\xi}{}^{s} = \zeta^{s}. \end{aligned}$$

Our results (9.4), (9.5) may thus be written in the form

$$\text{(9.7)}\quad \begin{aligned} &\text{(a)} \quad \underset{2}{\gamma_{00}} = 2\varphi = 2\psi + 2\chi, \\ &\text{(b)} \quad \underset{3}{\gamma_{0n}} = -2\psi\dot{\eta}^{n} - 2\chi\dot{\zeta}^{n}. \end{aligned}$$

From (1.18) we now also have

$$\text{(9.8)}\quad \begin{aligned} &\text{(a)} \quad \underset{2}{h_{00}} = \varphi = \psi + \chi, \\ &\text{(b)} \quad \underset{3}{h_{0n}} = -2\psi\dot{\eta}^{n} - 2\chi\dot{\zeta}^{n}, \\ &\text{(c)} \quad \underset{2}{h_{mn}} = \delta_{mn}\varphi = \delta_{mn}(\psi + \chi). \end{aligned}$$

This shows that the approximation $l = 1$ has a Newtonian character but, owing to the vanishing of $\underset{2}{\overset{k}{c}}{}_{m}$, places no restriction on the motion.

10. **Calculation of the Λ's for $l = 2$.** The first step in the calculation of the Λ's for $l = 2$ is the determination of the $h_{\mu\nu}$.

Using the method explained in 4, we can calculate the expansions of the $h^{\mu\nu}$ to any desired degree of approximation. We find, for $l = 1$,

$$\text{(10.1)}\qquad \underset{2}{h^{00}} = -\underset{2}{h_{00}} = -\varphi,$$

$$\text{(10.2)}\qquad \underset{3}{h^{0n}} = \underset{3}{h_{0n}} = \underset{3}{\gamma_{0n}},$$

$$\text{(10.3)}\qquad \underset{2}{h^{mn}} = -\underset{2}{h_{mn}} = -\delta_{mn}\varphi.$$

We next have to calculate for $l = 2$ the quantities $2L_{\mu\nu}$ defined in (1.15), (1.16), (1.17).

In $2\underset{4}{L_{00}}$ the linear terms give

$$\varphi_{,00}\,.$$

Of the non-linear terms, only three can give a contribution. They are

$$-2\left\{\underset{2}{h^{sr}}\left[\underset{2}{00, r}\right]\right\}_{,s} = -\varphi_{,s}\varphi_{,s} \quad (\text{since } \varphi_{,ss} = 0),$$

$$-2\left[\underset{2}{00, s}\right]\left[\underset{2}{0s, 0}\right] = \tfrac{1}{2}\varphi_{,s}\varphi_{,s}\,,$$

$$-2\left[\underset{2}{00, r}\right]\left[\underset{2}{rs, s}\right] = \tfrac{3}{2}\varphi_{,s}\varphi_{,s}\,,$$

where $[rs, p]$ are Christoffel symbols. Thus

$$2\underset{4}{L}_{00} = \varphi_{,00} + \varphi_{,s}\varphi_{,s}\,. \tag{10.4}$$

Similar but rather more tiresome calculations lead to the further results

$$2\underset{5}{L}_{0n} = \varphi_{,s}\underset{3}{h}_{0s,n} - \varphi_{,sn}\underset{3}{h}_{0s} - 3\varphi_{,0}\varphi_{,n}\,, \tag{10.5}$$

$$2\underset{4}{L}_{mn} = -\underset{3}{h}_{0m,0n} - \underset{3}{h}_{0n,0m} + \delta_{mn}\varphi_{,00} - 2\varphi\varphi_{,mn} - \varphi_{,m}\varphi_{,n} - \delta_{mn}\varphi_{,s}\varphi_{,s}\,. \tag{10.6}$$

Therefore, by (1.30), $\cdots$, (1.35), we have

$$\begin{aligned} &\text{(a)} \quad \underset{4}{\gamma}_{00,ss} = 2\underset{4}{\Lambda}_{00} = -\tfrac{3}{2}\varphi_{,s}\varphi_{,s}\,, \\ &\text{(b)} \quad \underset{5}{\gamma}_{0n,ss} = 2\underset{5}{\Lambda}_{0n} = \varphi_{,s}\underset{3}{\gamma}_{0s,n} - \varphi_{,sn}\underset{3}{\gamma}_{0s} - 3\varphi_{,0}\varphi_{,n}\,, \\ &\text{(c)} \quad \underset{4}{\gamma}_{mn,ss} = 2\underset{4}{\Lambda}_{mn} = -\underset{3}{\gamma}_{0m,0n} - \underset{3}{\gamma}_{0n,0m} + 2\delta_{mn}\varphi_{,00} - 2\varphi\varphi_{,mn} - \varphi_{,m}\varphi_{,n} \\ &\qquad\qquad + \tfrac{3}{2}\delta_{mn}\varphi_{,s}\varphi_{,s}\,. \end{aligned} \tag{10.7}$$

As explained in 7, 8, these equations (10.7), together with the corresponding coördinate conditions

$$\begin{aligned} &\text{(a)} \quad \underset{4}{\gamma}_{00,\underset{1}{0}} - \underset{5}{\gamma}_{0n,n} = 0, \\ &\text{(b)} \quad \underset{4}{\gamma}_{mn,n} = -\left(\underset{4}{\overset{1}{c}}_m/\overset{1}{r}\right) - \left(\underset{4}{\overset{2}{c}}_m/\overset{2}{r}\right), \end{aligned} \tag{10.8}$$

are the equations which determine the field in the next approximation.

11. **The Newtonian Equations of Motion.** We must now evaluate the surface integrals

$$\underset{4}{\overset{k}{c}}_m(\tau) = \frac{1}{4\pi}\int^k 2\underset{4}{\Lambda}_{mn}\cos(\mathbf{n}\cdot\mathbf{N})\,dS, \qquad k = 1, 2. \tag{11.1}$$

According to the general theory of part I, these integrals will be independent of the particular shapes of the surfaces of integration since the divergences of their integrands must vanish on a consequence of the field equations belonging to the previous approximation. We shall show here by actual calculation that this is the case with the $2\underset{4}{\Lambda}_{mn}$ given in (10.7c).

Since φ and $\underset{3}{\gamma}_{0n}$ are harmonic functions, we have

$$2\underset{4}{\Lambda}_{mn,n} = -\underset{3}{\gamma}_{0n,0mn} + 2\varphi_{,00m}$$

which is zero, as can easily be seen from (9.3) and (9.7a).

In the actual calculation of the surface integrals we evaluate the separate contributions of the different terms in $2\underset{4}{\Lambda}_{mn}$. Since the value of a whole integral is independent of the shape of its surface of integration, by taking this surface to be of finite size and always a finite distance from its singularity, we see that the whole integral cannot be infinite. Now the individual terms of $2\underset{4}{\Lambda}_{mn}$ have not the property that their divergences vanish, and so we must fix the surfaces of integration quite definitely before we begin the calculations. It is most convenient to take definite, infinitesimally small spheres whose centers are at the singularities, but in this case infinities of the types

$$\lim_{r\to 0} \text{const.}/r^n, \qquad n \text{ a positive integer,}$$

can occur in the values of the partial integrals. Since these must cancel, however, in the final result, we may merely ignore them throughout the calculation of the surface integrals.

We shall consider the integral taken around the first singularity. Owing to the infinitesimal size of the surface of integration, the only terms which can give results different from zero or infinity are those of the order of $\left(1/\overset{1}{r}{}^2\right)$.

The first term in $2\underset{4}{\Lambda}_{mn}$ is $-\underset{3}{\gamma}_{0m,0n}$, and, by (8.7b), this may be written as

$$-\underset{3}{\gamma}_{0m,0n} = -2\psi_{,ns}\dot{\eta}^m\dot{\eta}^s + 2\psi_{,n}\ddot{\eta}^m - 2\chi_{,ns}\dot{\zeta}^m\dot{\zeta}^s + 2\chi_{,n}\ddot{\zeta}^m.$$

The only term we need consider is the second, and so we have

$$\begin{aligned}\frac{1}{4\pi}\int^1\left(-\underset{3}{\gamma}_{0m,0n}\right)\cos(\mathbf{n}\cdot\mathbf{N})\,dS &= \frac{1}{4\pi}\int^1 2\psi_{,n}\ddot{\eta}^m\cos(\mathbf{n}\cdot\mathbf{N})\,dS \\ &= \left(4\overset{1}{m}\ddot{\eta}^m\right)\frac{1}{4\pi}\int^1\left\{(x^n-\eta^n)(x^n-\eta^n)/\overset{1}{r}{}^4\right\}dS \qquad (11.2)\\ &= \left(4\overset{1}{m}\ddot{\eta}^m\right)\frac{1}{4\pi}\int^1\left\{1/\overset{1}{r}{}^2\right\}dS = 4\overset{1}{m}\ddot{\eta}^m.\end{aligned}$$

In a similar manner we find that

$$\frac{1}{4\pi}\int^1\left(-\underset{3}{\gamma}_{0n,0m}\right)\cos(\mathbf{n}\cdot\mathbf{N})\,dS = \tfrac{4}{3}\overset{1}{m}\ddot{\eta}^m. \qquad (11.3)$$

The fourth term, $(-2\varphi,\ \varphi_{,mn})$, requires slightly different treatment. The only part that can be of interest is

$$-2\psi_{,mn}\,\chi$$

and in order to evaluate the corresponding contribution to the surface integral we must expand χ as a power series in the neighborhood of the first singularity, writing

$$\chi = \tilde{\chi} + (x^s - \eta^s)\tilde{\chi}_{,s} + \cdots, \qquad (11.4)$$

where

$$\tilde{\chi} = \chi(\eta^n), \qquad \tilde{\chi}_{,s} = \chi_{,s}(\eta^n), \text{ etc.} \qquad (11.5)$$

Introducing this expansion for χ we see that the only term in the integrand which can give a finite result is

$$-2\psi_{,mn}(x^s - \eta^s)\tilde{\chi}_{,s} \tag{11.6}$$

The determination of the surface integral of this term depends on the calculation of

$$\int^1 (x^s - \eta^s)\, \psi_{,mn} \cos(\mathbf{n}\cdot\mathbf{N})\, dS.$$

We have

$$\begin{aligned}(x^s - \eta^s)\, \psi_{,mn} \cos(\mathbf{n}/\mathbf{N}) \\ &= 2\overset{1}{m}(x^s - \eta^s)\left\{-3(x^m - \eta^m)(x^n - \eta^n)/\overset{1}{r}{}^5 + \delta_{mn}/\overset{1}{r}{}^3\right\}(x^n - \eta^n)/\overset{1}{r} \\ &= -4\overset{1}{m}(x^s - \eta^s)(x^m - \eta^m)/\overset{1}{r}{}^4.\end{aligned}$$

Therefore

$$\begin{aligned}\frac{1}{4\pi}\int (x^s - \eta^s)\, \psi_{,mn} \cos(\mathbf{n}\cdot\mathbf{N})\, dS &= -\frac{4\overset{1}{m}}{4\pi}\int (x^s - \eta^s)(x^m - \eta^m)/\overset{1}{r}{}^4\, dS \\ &= -\frac{4\overset{1}{m}}{3}\,\delta_{ms},\end{aligned} \tag{11.7}$$

and so the surface integral of the term (11.6) is

$$+\frac{8\overset{1}{m}}{3}\,\tilde{\chi}_{,m}, \tag{11.8}$$

which is thus also the value of the surface integral for the whole of the term $(-\varphi\varphi_{,mn})$.

In a somewhat similar way we obtain, for the surface integrals of the remaining terms, the values

$$\left\{\begin{aligned} 2\delta_{mn}\varphi_{,00} &\to -\frac{4\overset{1}{m}}{3}\,\ddot{\eta}^m, \\ -\varphi_{,m}\varphi_{,n} &\to -\frac{8\overset{1}{m}}{3}\,\tilde{\chi}_{,m}, \\ \tfrac{3}{2}\delta_{mn}\varphi_{,s}\varphi_{,s} &\to 2\overset{1}{m}\tilde{\chi}_{,m}. \end{aligned}\right. \tag{11.9}$$

Hence we have

$$\underset{4}{\overset{1}{c}}{}_m(\tau) = \frac{1}{4\pi}\int^1 2\underset{4}{\Lambda}{}_{mn} \cos(\mathbf{n}\cdot\mathbf{N})\, dS = 4\overset{1}{m}\{\ddot{\eta}^m + \tfrac{1}{2}\tilde{\chi}_{,m}\}. \tag{11.10}$$

Let us assume for the moment that we are not going any further with the approximation. In this case our approximate equations of motion would be of the form

$$\lambda^4\{\ddot{\eta}^m + \tfrac{1}{2}\tilde{\chi}_{,m}\} = 0 \tag{11.11}$$

for each particle. It is of interest to note that this form of the equations of motion is actually independent of the variables x^s. For we have, by (11.5), (9.6),

$$\tilde{\chi}_{,s} = \chi_{,s}(\eta^n), \qquad \chi = -2\overset{2}{m}/\overset{2}{r}. \tag{11.12}$$

For our present argument we may take χ as any function of $\overset{2}{r}$. Equations (11.12) show that to form $\tilde{\chi}_{,s}$ we must first differentiate χ with respect to x^s and then replace x^s by η^s. But the result will be the same if we first replace x^s by η^s and later differentiate with respect either to η^s or to $(-\zeta^s)$. Thus

$$\tilde{\chi}_{,s} = \frac{\partial\chi(r)}{\partial\eta^s} = -\frac{\partial\chi(r)}{\partial\zeta^s}, \tag{11.13}$$

where r denotes the "distance" between η^s and ζ^s:

$$r = [(\eta^s - \zeta^s)(\eta^s - \zeta^s)]^{\frac{1}{2}}. \tag{11.14}$$

We can therefore think of our equations of motion as involving the differentiation of functions depending only on the positions of the singularities, as is characteristic of theories based on the concept of action at a distance.

Writing (11.11) more explicitly in vector notation as

$$\overset{1}{m}\ddot{\eta} = \nabla\left(\overset{1}{m}\overset{2}{m}/r\right), \tag{11.15}$$

we see that (11.11) gives precisely the Newtonian law of motion.[11]

We have therefore obtained the Newtonian equations of motion from the field equations alone, without extra assumption such as was hitherto believed to be necessary and was supplied by the law of geodetic lines, or by a special choice of an energy impulse tensor.

From the above derivation of the Newtonian equations of motion, the general mechanism becomes apparent by which the Lorentz equations for the motion of electric particles can be obtained. In this case we have to consider the gravitational equations in which the Maxwell energy-momentum tensor appears on the right, and also the Maxwell field equations, and treat the whole set of equations by our approximation method. It is necessary, now, to give each singularity an electric charge e in addition to its mass m. We may safely ignore the contribution arising from the products of gravitational potentials in

[11] Equation (11.11) and (11.15) are written in terms of the auxiliary time and the auxiliary masses. We shall return to this point in 17.

the new field equations. For this omission has the effect of destroying the second term of (11.11), while the inclusion of the Maxwell tensor leads to the appearance, on the right of (11.11), of a corresponding surface integral giving the electrostatic force acting on the particle. In the next approximation we obtain the full Lorentz force together with the relativistic correction to the mass.

So long as we are dealing with singularities, we have no basis within the theory for excluding negative masses; in other words, for excluding gravitational repulsions between particles. If, however, we decide always to take mass positive, then the sign with which the Maxwell energy-momentum tensor enters the field equations determines whether like charges shall attract or repel each other. This also reveals the limitations of any theory based upon the existence of singularities.

12. **Normalisation of $\underset{4}{\gamma}_{00}$.** The value of $\underset{4}{\gamma}_{00}$ determined from (10.7a) is arbitrary to within an added harmonic function, and this function is to be determined from the relations (8.4), (8.2), together with our basic requirement that higher harmonic functions than simple poles are, as far as possible, to be avoided.

From (10.7a) and the fact that φ is harmonic, we have at once

$$\underset{4}{\gamma}_{00} = -\tfrac{3}{4}\varphi\varphi + \alpha_{00}\psi + \beta_{00}\chi, \tag{12.1}$$

where we have written the additive functions of (8.4) in a different form more in accordance with our present notation, α_{00}, β_{00} being functions of τ alone through η and ζ and their derivatives. The quantities α_{00}, β_{00} can be determined from the condition that

$$\frac{1}{4\pi}\int\left\{\underset{4}{\gamma}_{00,0n} - 2\underset{5}{\Lambda}_{0n}\right\}\cos(\mathbf{n}\cdot\mathbf{N})\,dS = 0. \tag{12.2}$$

The value of α_{00} is found by taking this integral over a small sphere having its center at the first singularity, and from calculations similar to those of 3, we find, after making use of the equations of motion of the first order:

$$\alpha_{00} = \{\dot{\eta}^s\dot{\eta}^s + \tfrac{1}{2}\tilde{\chi}\}. \tag{12.3}$$

Similarly, by integrating over a small sphere around the second singularity, we find

$$\beta_{00} = \{\dot{\zeta}^s\dot{\zeta}^s + \tfrac{1}{2}\tilde{\psi}\}, \tag{12.4}$$

where

$$\begin{aligned}\tilde{\chi} &= \chi(\eta^s),\\ \tilde{\psi} &= \psi(\zeta^s).\end{aligned} \tag{12.5}$$

These results show clearly the physical significance of the particular normalisation required by the conditions (8.4), (8.2). For we now have

$$\lambda^2 \underset{2}{\gamma_{00}} + \lambda^4 \underset{4}{\gamma_{00}} = \lambda^2 \{ (1 + \tfrac{1}{2}\lambda^2 \alpha_{00})\, 2\psi + (1 + \tfrac{1}{2}\lambda^2 \beta_{00})\, 2\chi - \tfrac{3}{4}\lambda^2 \varphi\varphi \}, \tag{12.6}$$

and we see from (12.3), (12.4) that $\overset{1}{m}(1 + \frac{1}{2}\lambda^2 \alpha_{00})$, $m(1 + \frac{1}{2}\lambda^2 \beta_{00})$ involve the first relativistic corrections to the masses.

The calculations up to this stage correspond to those of Droste, De Sitter, and Levi-Civita, cited in the introduction.

13. Solution of the Field Equations for $l = 2$. Since our ultimate aim is to determine the equations of motion up to the next approximation, we are interested only in those expressions which give a contribution to the corresponding surface integrals. We shall state dogmatically what is needed for these calculations for the justification of our statement can not be given without exposing the details of our actual calculation.

1. The calculation of $\underset{4}{\gamma_{mn}}$ and $\underset{5}{\gamma_{0m}}$ in the neighborhood of the singularities. We do not need to care in $\underset{4}{\gamma_{mn}}$ about those terms which do not go to infinity if $\overset{1}{r} \to 0$.

2. The calculation of $\underset{4}{\gamma_{rr}}$ in the whole space.

The expression $2\underset{4}{\Lambda_{mn}}$ in (10, 7) can be divided into two parts, one containing the linear terms together with all other terms not involving interactions between the two particles, and the other containing all the interaction terms. We denote these two groups of terms respectively by X_{mn} and Y_{mn}. The integration of the equations

$$\underset{4}{\gamma'_{mn,ss}} = X_{mn} \tag{13.1}$$

presents no difficulties, but the equations

$$\underset{4}{\gamma''_{mn,ss}} = Y_{mn} \tag{13.2}$$

cannot, apparently, be integrated in an elementary manner and we are obliged to introduce a simplification. Since we need to know the values of $\underset{4}{\gamma_{mn}}$ mainly in order to evaluate the surface integrals $\underset{6}{c_m}$ about, say, the first particle, we may introduce power series expansions for χ in the neighborhood of this point and so obtain a solution for $\underset{4}{\gamma_{mn}}$ which is also in the form of such an expansion.

We find, actually, from (13.1), (13.2), the following expressions for $\underset{4}{\gamma}'_{mn}$ and $\underset{4}{\gamma}''_{mn}$:

$$\begin{aligned}\underset{4}{\gamma}'_{mn} = {} & \{\psi[(x^n - \eta^n)\dot{\eta}^m + (x^m - \eta^m)\dot{\eta}^n - \delta_{mn}(x^s - \eta^s)\dot{\eta}^s]\}_{,0} \\ & + \{\chi[(x^n - \zeta^n)\dot{\zeta}^m + (x^m - \zeta^m)\dot{\zeta}^n - \delta_{mn}(x^s - \zeta^s)\dot{\zeta}^s]\}_{,0} \\ & + \tfrac{7}{4}\overset{1}{r}{}^2\psi_{,m}\psi_{,n} + \tfrac{7}{4}\overset{2}{r}{}^2\chi_{,m}\chi_{n},\end{aligned} \tag{13.3}$$

and

$$\underset{4}{\gamma}''_{mn} = -\psi_{,m}(x^n - \eta^n)\tilde{\chi}, \tag{13.4}$$

where we have included in (13.4) only those terms which ultimately have importance for the evaluation of the surface integrals $\underset{6}{\overset{1}{c}}_m$.

The value of $\underset{4}{\gamma}_{mn}$ is given by

$$\underset{4}{\gamma}_{mn} = \underset{4}{\gamma}'_{mn} + \underset{4}{\gamma}''_{mn} + \alpha_{mn}\psi, \tag{13.5}$$

where α_{mn} is a function of time to be determined from the coördinate conditions

In a similar way, we may calculate the values of $\underset{5}{\gamma}_{0n}$ in two parts. We find, on including only relevant terms for the surface integrals $\underset{6}{\overset{1}{c}}_m$, in the integrands of which $\underset{5}{\gamma}_{0n}$ enters only linearly,

$$\underset{5}{\gamma}'_{0m} = -\tfrac{7}{4}\overset{1}{r}{}^2\psi_{,m}\psi_{,s}\dot{\eta}^s + \tfrac{3}{4}\psi\psi\dot{\eta}^m, \tag{13.6}$$

$$\begin{aligned}\underset{5}{\gamma}''_{0n} = {} & -\tfrac{3}{2}(x^s - \eta^s)\psi\tilde{\chi}_{,m}\dot{\zeta}^s - (x^m - \eta^m)\psi\tilde{\chi}_{,s}\dot{\eta}^s \\ & + \tfrac{1}{2}\psi_{,m}(x^s - \eta^s)(x^l - \eta^l)\tilde{\chi}_{,l}\dot{\zeta}^s + (x^m - \eta^m)\psi\tilde{\chi}_{,s}\dot{\zeta}^s \\ & + \tfrac{1}{2}(x^s - \eta^s)\psi\tilde{\chi}_{,s}\dot{\zeta}^m + \tfrac{3}{2}(x^s - \eta^s)\psi\tilde{\chi}_{,m}\dot{\eta}^s \\ & + (x^s - \eta^s)\psi_{,m}\tilde{\chi}\dot{\zeta}^s.\end{aligned} \tag{13.7}$$

The value of $\underset{5}{\gamma}_{0n}$ is given by

$$\underset{5}{\gamma}_{0n} = \underset{5}{\gamma}'_{0n} + \underset{5}{\gamma}''_{0n} + \alpha_{0n}\psi, \tag{13.8}$$

where α_{0n} is a function of time to be determined from the normalisation condition.

It remains only to calculate $\underset{4}{\gamma}_{rr}$ in the whole space. From (10.7c) we have

$$\underset{4}{\gamma}_{rr,ss} = 2\varphi_{,00} + \tfrac{7}{2}\varphi_{,s}\varphi_{,s}, \tag{13.9}$$

therefore

$$\underset{4}{\gamma}_{rr} = -2\overset{1}{m}\overset{1}{r}_{,00} - 2\overset{2}{m}\overset{2}{r}_{,00} + \tfrac{7}{4}\varphi^2 + \alpha\psi + \beta\chi \tag{13.10}$$

where α and β are functions of time to be determined in such a way that $\underset{4}{\gamma}_{rr}$ in (13.10) would agree with $\underset{4}{\gamma}_{rr}$ determined from (13.5) near to the singularities.

14. **Determination of α_{mn} and α_{0n}.** In order to find α_{mn}, α_{0n} from the conditions (8.7e), (8.7c) we must make use of the values of $\underset{4}{\overset{1}{c}}_m$ found in 3. The result is up to the desired order

$$\alpha_{mn} = \{2\dot{\eta}^m\dot{\eta}^n + \delta_{mn}\tilde{\chi}\} \tag{14.1}$$

and

$$\alpha_{0n} = -\dot{\eta}^s\dot{\eta}^s\dot{\eta}^n + \tilde{\chi}\dot{\eta}^n - \tilde{\chi}\dot{\zeta}^n. \tag{14.2}$$

Finally from our last remark in 13 follows:

$$\alpha = 2\dot{\eta}^s\dot{\eta}^s + \tfrac{1}{2}\tilde{\chi}; \qquad \beta = 2\dot{\zeta}^s\dot{\zeta}^s + \tfrac{1}{2}\tilde{\psi}. \tag{14.3}$$

15. **Calculation of $\underset{6}{\Lambda}_{mn}$.** In the calculation of $\underset{6}{\Lambda}_{mn}$ for our present purposes, we may assume that $\underset{4}{c}_m$ is zero, as we shall now show.

After we have evaluated the surface integrals $\underset{6}{c}_m$, we may write the approximate equations of motion in the form

$$\lambda^4\underset{4}{c}_m + \lambda^6\underset{6}{c}_m = 0. \tag{15.1}$$

But this shows that when the motion is in accordance with (15.1) the quantities $\lambda^4\underset{4}{c}_m$ and $\lambda^6\underset{6}{c}_m$ will be of the same order of magnitude. It is evident, however, that $\lambda^4\underset{4}{c}_m$ can enter $\lambda^6\underset{6}{\Lambda}_{mn}$ only in combination with a quantity of the type $\lambda^2\underset{2}{\Theta}$. It will therefore enter only in terms which actually belong to the order λ^8 or higher, and since we do not propose to go beyond the order λ^6 in the calculation of the equations of motion, we may neglect all terms in $\underset{6}{\Lambda}_{mn}$ in which $\underset{4}{c}_m$ appears. Even if we make use of this fact, however, the calculations are still quite tedious, and there are actually forty-one different types of term in the expansion of $\underset{6}{\Lambda}_{mn}$. We find:

$$
\begin{aligned}
2\underset{6}{\Lambda}_{mn} = {} & -\underset{5}{\gamma}_{0m,0n} - \underset{5}{\gamma}_{0n,0m} + \delta_{mn}\underset{4}{\gamma}_{00,00} + \underset{4}{\gamma}_{mn,00} - \varphi\underset{4}{\gamma}_{00,mn} - \varphi\underset{4}{\gamma}_{ss,mn} - \varphi_{,mn}\underset{4}{\gamma}_{00} \\
& - \varphi_{,mn}\underset{4}{\gamma}_{ss} + \varphi_{,ms}\underset{4}{\gamma}_{ns} + \varphi_{,ns}\underset{4}{\gamma}_{ms} - \delta_{mn}\varphi_{,sr}\underset{4}{\gamma}_{sr} - 2\varphi_{,s}\underset{4}{\gamma}_{mn,s} + \varphi_{,s}\underset{4}{\gamma}_{ms,n} \\
& + \varphi_{,s}\underset{4}{\gamma}_{ns,m} - \tfrac{1}{2}\varphi_{,m}\underset{4}{\gamma}_{ss,n} - \tfrac{1}{2}\varphi_{,n}\underset{4}{\gamma}_{ss,m} - \tfrac{1}{2}\varphi_{,n}\underset{4}{\gamma}_{00,m} - \tfrac{1}{2}\varphi_{,m}\underset{4}{\gamma}_{00,n} \\
& + \tfrac{3}{2}\delta_{mn}\varphi_{,s}\underset{4}{\gamma}_{rr,s} + \tfrac{3}{2}\delta_{mn}\varphi_{,s}\underset{4}{\gamma}_{00,s} - \underset{3}{\gamma}_{0s}\underset{3}{\gamma}_{0n,ms} - \underset{3}{\gamma}_{0s}\underset{3}{\gamma}_{0m,ns} + 2\underset{3}{\gamma}_{0s}\underset{3}{\gamma}_{0s,mn} \\
& + \tfrac{1}{2}\delta_{mn}\underset{3}{\gamma}_{0s,r}\underset{3}{\gamma}_{0r,s} - \tfrac{3}{2}\delta_{mn}\underset{3}{\gamma}_{0s,r}\underset{3}{\gamma}_{0s,r} + \underset{3}{\gamma}_{0s,m}\underset{3}{\gamma}_{0s,n} + \underset{3}{\gamma}_{0m,s}\underset{3}{\gamma}_{0ns} - \varphi_{,0n}\underset{3}{\gamma}_{0m} \\
& - \varphi_{,0m}\underset{3}{\gamma}_{0n} + 2\delta_{mn}\varphi_{,0s}\underset{3}{\gamma}_{0s} - \varphi_{,0}\underset{3}{\gamma}_{0m,n} - \varphi_{,0}\underset{3}{\gamma}_{0n,m} - \varphi_{,n}\underset{3}{\gamma}_{0m,0} - \varphi_{,m}\underset{3}{\gamma}_{0n,0} \\
& + 2\varphi\underset{3}{\gamma}_{0m,0n} + 2\varphi\underset{3}{\gamma}_{0n,0m} - 2\delta_{mn}\varphi\varphi_{,00} + 2\varphi\varphi\varphi_{,mn} - \varphi\varphi_{,m}\varphi_{,n} \\
& + \tfrac{3}{2}\delta_{mn}\varphi\varphi_{,s}\varphi_{,s} + \tfrac{1}{2}\delta_{mn}\varphi_{,0}\varphi_{,0}\,.
\end{aligned}
\tag{15.2}
$$

The condition that $\underset{6}{\Lambda}_{mn,n}$ must be zero affords a valuable test of the correctness of the above formula. We have worked out the divergence of the $\underset{6}{\Lambda}_{mn}$ given in (15.2) and have found that it does indeed vanish.

16. **The Surface integrals for $l = 3$.** In order to find the principal deviation from the Newtonian laws of motion, all that essentially remains is to calculate the values of the surface integrals $\underset{6}{c}_m$. To do this we must first insert in (15.2) the values previously found for $\underset{4}{\gamma}_{00}$, $\underset{4}{\gamma}_{mn}$ and $\underset{5}{\gamma}_{0n}$ and then it is a matter of calculating the contributions of the resulting terms one by one and adding the expressions obtained. The general technique is similar to that used in 11 for the evaluation of $\underset{4}{c}_m$ but considerably more complicated.

On making use of our right to take $\underset{4}{c}_m$ to be zero, we may express the result in the form

$$
\begin{aligned}
\overset{1}{\underset{6}{c}}_m &= \frac{1}{4\pi}\int^{1} 2\underset{6}{\Lambda}_{mn}\cos(\mathbf{n}\cdot\mathbf{N})\,dS \\
&= -4\overset{1}{m}\overset{2}{m}\Bigg\{\left[\dot\eta^s\dot\eta^s + \tfrac{3}{2}\dot\zeta^s\dot\zeta^s - 4\dot\eta^s\dot\zeta^s - 4\frac{\overset{2}{m}}{r} - 5\frac{\overset{1}{m}}{r}\right]\frac{\partial}{\partial\eta^m}\left(\frac{1}{r}\right) \\
&\qquad [4\dot\eta^s(\dot\zeta^m - \dot\eta^m) + 3\dot\eta^m\dot\zeta^s - 4\dot\zeta^s\dot\zeta^m]\frac{\partial}{\partial\eta^s}\left(\frac{1}{r}\right) + \frac{1}{2}\,\frac{\partial^3 r}{\partial\eta^s\,\partial\eta^r\,\partial\eta^m}\dot\zeta^s\dot\zeta^r\Bigg\}.
\end{aligned}
\tag{16.1}
$$

17. The Main Deviation from the Newtonian Equations of Motion. In order to obtain the equations of motion belonging to this stage of our approximation, we must write

$$\lambda^4 \overset{k}{\underset{4}{c}}_m + \lambda^6 \overset{k}{\underset{6}{c}}_m = 0 \qquad k = 1, 2 \tag{17.1}$$

and then must reabsorb the λ's by going over to the old time x^0 instead of the auxiliary time $\tau = \lambda x^0$ and by introducing a corresponding change in mass from m to M, where $M = \lambda^2 m$. There will be no confusion if we keep the old notation for the new quantities so that now $\dot{\xi} = d\xi/dx^0$ instead of $d\xi/d\tau$, and m is written for the new mass M. And with this convention we may write the equations of motion (17.1), by means of (11.10) and (16.1), in the form

$$\begin{aligned} \ddot{\eta}^m - \overset{2}{m} \frac{\partial(1/r)}{\partial \eta^m} = \overset{2}{m} \Bigg\{ & \left[\dot{\eta}^s \dot{\eta}^s + \tfrac{3}{2} \dot{\zeta}^s \dot{\zeta}^s - 4 \dot{\eta}^s \dot{\zeta}^s - 4 \frac{\overset{2}{m}}{r} - 5 \frac{\overset{1}{m}}{r} \right] \frac{\partial}{\partial \eta^m} (1/r) \\ & [4\dot{\eta}^s(\dot{\zeta}^m - \dot{\eta}^m) + 3\dot{\eta}^m \dot{\zeta}^s - 4\dot{\zeta}^s \dot{\zeta}^m] \frac{\partial}{\partial \eta^s} (1/r) + \frac{1}{2} \frac{\partial^3 r}{\partial \eta^s \partial \eta^r \partial \eta^m} \dot{\zeta}^s \dot{\zeta}^r \Bigg\}. \end{aligned} \tag{17.2}$$

The equations of motion for the other particle are obtained by replacing $\overset{1}{m}$, $\overset{2}{m}$, η, ζ by $\overset{2}{m}$, $\overset{1}{m}$, ζ, η.

These equations, giving the relativistic motion of two massive gravitating bodies, constitute the main result of our calculations from the point of view of practical application.

These equations have since been integrated by H. P. Robertson, whose results are given in the following note on "The Two Body Problem in General Relativity," Math. Ann. 39, p. 101 (1938).

We should like to thank Professor Robertson for the very kind interest he took in this problem and for his help.

THE INSTITUTE FOR ADVANCED STUDY.

Chapter 16
Erwin Schrödinger (1940): The General Theory of Relativity and Wave Mechanics

Erwin Schrödinger (1940). The General Theory of Relativity and Wave Mechanics. *Wis- en natuurkundig Tijdschrift*, 10: 2–9.

The General Theory of Relativity and Wave Mechanics.

By ERWIN SCHRÖDINGER (at present Dublin).

1. *Curved space-time — the model of matter.* — The most important discoveries are those which in the course of time tend to become tautological. The logical content of Newton's first two laws of motion was to state, that a body moves uniformly in a straight line, unless it does something else and that in the latter case we agree upon *calling force* its acceleration multiplied by an individual constant. The great achievement was, to concentrate attention on the *second* derivatives — to suggest that *they* — not the first or third or fourth, not any other property of the motion — ought to be accounted for by the environment. It is very often more essential just to point out the kind of conception that matters and to emphasize it by claiming for it a name of its own, than to find out in detail the laws which control it.

The fundamental statements of Einstein's theory of gravitation are of a similar kind. The equations

$$G_{ik} - \tfrac{1}{2} g_{ik} G = -8\pi T_{ik} \tag{1}$$

state, that the contracted curvature tensor G_{ik} is either zero or not (1) and that, when and where it is not, we *call matter* (T_{ik}) the left hand side of equations (1). Thus the equations amount to a definition of *matter* (T_{ik}), on which they impose the *conservation laws*, in virtue of a well known identity, satisfied by the G_{ik} of *any* metric g_{ik}. *They do not impose any restriction whatsoever on the metric* (g_{ik}). Yet this definition is all-important. Matter is no longer the content of the receptacle space-time, is no longer the actor who performs in time on the stage of space. Matter and space-time have really merged into one. It is sometimes said, that matter determines the curvature of space-time. But the most advisable attitude is, I think, to reserve the expression of *matter* to indicate the object of our direct observation and to regard *curved space-time* as the picture or *model* we form of this object in our minds, bearing thus the same relation to matter as the Rutherford-Bohr model to the atom of observation. Equation (1) enounces the choice of the model. Our observations on matter determine *us* to equip the model we have chosen with such features (curvature) as make it fit in with observation.

---

(1) See A. S. Eddington, The mathematical theory of relativity, Ch. IV, sect. 54.

2. *The model gives all forces the same standing.* — Though I do not expect to be opposed in what I just said, I wish to point out that it compels us to change the wording of some familiar statements. It is very often said that on the atomic scale and even on the macroscopic scale no influence of gravitation is perceptible, apart from the intense gravitational fields of the huge celestial bodies. Now look at that book lying on your desk. Why does it not fall in spite of the gravitational pull of the planet? On account of the elastic pressure, equal and of opposite sign, which the table exerts on it. Of what kind is that pressure? You may begin to tell me of deformed cristal-lattices, deformed electronic states, particle-encounters, exchange-forces, etc. To which I reply, that no doubt you are right. But whether your details are correct or not, the result of all that is certainly, that some T_{ik}, very well known from the theory of elasticity, are aroused in the table and in the book. These T_{ik} form part of the right hand side of equation (1) and compel us to equip our space-time continuum whith the corresponding curvature G_{ik}. And that is essential for the book being maintained in its place in our model. The circumstances which would cause the book to fall, if the table were not there, and those which actually prevent it from falling are thus of exactly the same kind, they have exactly the same standing. If it is said, as it often is, that the essential thing about Einstein's theory is to explain gravitation by the curvature of space-time, then one has to declare, that all forces — i. e. all that was formerly called force — are gravitational. Gravitation in this sense is not to be coordinated with forces « of other origin », it is the comprehensive conception of them all.

Take an electron, moving in an electrostatic field in, otherwise, empty space. Why does it not follow the geodesic? The usual answer is: because it is charged and there is the electric pull on it. The correct answer is: 1) Only in empty space ($T_{ik} = 0$) is a particle, under certain assumptions, expected to follow the geodesic. Our electron is moving through matter ($T_{ik} \neq 0$). 2) It is true, that other (sc. neutral) particles would under the same circumstances follow the geodesic; that our electron does not, is due to certain modifications (additional T_{ik}) which it produces by reacting with the matter through which it moves. The case resembles that of a gun-ball moving in air, which, for similar reasons, does not follow the parabola.

Whereas the gravitational field which in the popular sense of the word is produced by the energy or mass (T_{44}) of an electromagnetic field is practically negligible, we see from the preceeding « example », that the curvature corresponding to components like T_{11}, T_{23} etc. is appreciable and rather efficient. That is partly due to the fact, that in the case of an electromagnetic field these components are of the same order of magnitude as T_{44} (which in ordinary matter they are usually not); partly to the particularly favourable spatial distribution of those T_{ik} which are aroused by the interaction of the two electric fields, that of the electron and that through which it moves.

3. *The model does not determine particle motion.* — It was emphasized in § 1, that the equations (1) do not impose any restriction on the metric (g_{ik}). Let us choose an arbitrary metric (g_{ik}) and calculate its matter (T_{ik}) from (1). Let us assume, that there are empty regions, i. e. with $T_{ik} = 0$. Now choose, quite arbitrarily, a four-dimensional track, which passes also through empty regions. Choose among the *neighbouring* metrics $g_{ik} + \delta g_{ik}$ one for which the $T_{ik} + \delta T_{ik}$, calculated from (1), differ from the T_{ik} only in the vicinity of the chosen track. That is certainly possible, for you might restrict the δg_{ik} themselves to the vicinity of the track (though that would be a very special choice).

In those parts where the track passes through regions that were empty in the metric g_{ik} it will, in the metric $g_{ik} + \delta g_{ik}$, have certain features of the world line of a moving particle. But since the track was arbitrarily chosen, it need not, of course, be a geodesic of the original metric. It may include space-like directions, may (by passing through such) change from positive to negative time-like directions, may begin or end suddenly in previously empty space-time, may exhibit any desired kind of bifurcations etc.

That the field equations (1) have no tendency whatsoever of guiding a small disturbance in vacuo along a geodesic, is a well known fact (1). (I have insisted on recalling it here, because it is liable to be veiled again and again by the treatment in some text books.) What then do they tell us about this case ? Nothing beyond the conservation identities. For an isolated disturbance in vacuo that means *the parallel displacement of the resultant energy momentum-vector along the path* — provided that there is a path, i. e. as long as the disturbance keeps together.

Neither do the field equations take charge of keeping the disturbance together — they could not, for they have to be capable of describing events, where it is not the case. Nor do they say anything about the track, because any four-vector can be parallelly displaced along any track. One thing can be said : if you choose to let the path begin or end in vacuo, you may be sure to find the energy-momentum vector zero at every cross section. So this case would necessarily introduce

(1) A. S. EDDINGTON, The mathematical Theory of Relativity, Ch, 4, sect. 56.

negative mass density. A bifurcation of such a track of total mass zero into a positive and a negative branch would be perfectly admitted, and so would the confluence of two such tracks be, e.g. with subsequent annihilation.

In order to preclude these and other odd possibilities, contrary to observation, current theory introduces the following assumptions : 1) mass is always positive, 2) it is concentrated in particles that are ordinarily indivisible, 3) the (four dimensional) tangent of the track continues to point in the direction of the resultant energy-momentum vector. — The third assumption, when combined with the conservation law, introduces the *geodesic* track rather openly as a distinct hypothesis. (It may be pointed out, though, that it would be pretty difficult to introduce, in an equally simple way, *another* law for the track.)

4. *There is room for wave-mechanics to supplement the model.* — These points have, of course, always formed an integrant part of the theory, so much so, that one is liable to take them for granted and to overlook the fact, that the field equations need any supplement at all. Indeed the latter was an almost trivial and straightforward loan from the particle-picture of matter, which was well developed at that time, whereas the discovery of the field equations was the achievement of unusual geniality and painstaking.

There can be little doubt, that from the outlook we have reached today we have to convert this loan by putting wave-mechanics in the place of particle mechanics. What I wish to emphasize, is *that there is room and even need for supplementing the general theory of relativity by wave-mechanics*. You could not expect relativity theory to produce out of its own something of that kind, say e.g. in the way that gravitational waves turned out from equations (1) to follow laws similar to those of material waves. If you wish them to do that, you have to introduce it as an explicit assumption, and you are allowed to do it. For the field equations themselves have just as little control of the behaviour of a wave-like disturbance as of that of a particle-like disturbance.

Wave-mechanics seems to be very fit for furnishing the wanted. Anyone of the current wave-equations, when placed in an arbitrary Riemannian space and disencumbered of the terms corresponding to explicit forces, will — in virtue of the Fermat-Hamilton-principle — describe a wave propagation along the geodesics of that space. Moreover in the case of parcel-like solutions, which are always available, the theory affords a definition of the momentum vector, which actually undergoes, from the very laws of the wave motion, a parallel displacement *in its own direction*. It is true, that things are not quite as simple as I have made them look by these statements — inter alia, because they primarily refer to threedimensional space. But a generalisation to include time is quite possible.

One has the impression, that wave-mechanics meets the requirements of general relativity even in a more congenial way than particle mechanics did. That a particle-like solution does not, as a rule, keep together for ever and that, moreover, solutions of entirely different type exist, need not worry us, because we know what it means (principle of indeterminacy). Further we know, that the demarcation of individuals, corresponding to the observed discontinuity of matter, finds an entirely different expression in wave-mechanics, not by delimiting the individuals in space, but by the discreteness in energy-momentum space (line-spectrum).

5. *Two geometrical aspects of matter : the differential and the integral one.* — From our present outlook on Nature there is, I think, no other way of accounting for the atomicity of matter than by admitting the eigenvalues of the wave-equation to be discrete, to form a line-spectrum. And there is, I believe, no other general means of having them so, than by assuming space to be closed and finite. If these conclusions are accepted, they open far reaching prospects.

The proper modes of a given closed space, since they do not depend on boundary conditions, may be regarded as a purely geometrical property of that space. The wave aspect connects matter with these proper modes, thus with pure geometry — as the space-time model of general relativity also does. But the two connexions are of entirely different kind. The one is a local and differential one : matter here and now is described by the curvature here and now. The other is an integral connexion. The very nature of the matter which forms the Universe is determined by the shape and size of the Universe as a whole. I believe that these two aspects are not contradictory but complementary.

There is an interesting coincidence to confirm these views. All observed masses are positive. Positive mass has, to say the least, a marked tendency of producing a positive curvature of *space* — such is, at any rate, the current interpretation of the state of affairs within our galaxy. A continuum that in all its parts has on the average a positive curvature, cannot, I think, avoid singularities without being closed. But from the uncertainty principle we may conclude, that a closed configuration space, since it sets an upper limit to the uncertainty of position, entails a positive kinetic zero-point energy, which gives rise to positive mass. Thus closed space is, in a way, self-consistent. There is a remarkable general agreement between the two geometrical connexions in that they both suggest a positive density of matter — as is also observed.

6. *Attempts to account for the elementary mass.* — The ideas I have just expounded are due to A. S. Eddington [1]. We ought to discriminate carefully between them and Eddington's elaborate attempt to carry them out. Let me in the simplest and most primitive terms (for which I alone am responsible) outline the difficulties met with.

A first naive trial would run thus. If R is the radius of a spherical Universe, configuration space is restricted to the linear dimension R. The momentum p of the lowest state will therefore be of the order of magnitude

$$p = \frac{h}{2\pi R} \cdot \tag{2}$$

We associate with p an energy $\frac{p^2}{2\,m}$ and put it equal to the rest-or zero-point-energy $m\,c^2$. Thus

$$\frac{1}{2m}\left(\frac{h}{2\pi R}\right)^2 = mc^2$$

or

$$m = \frac{h}{2\pi Rc\sqrt{2}} \cdot \tag{3}$$

To give the mass of the proton, the radius of the Universe, R, would have to be of the order $h/mc \approx 10^{-13}$ cm, wich is pure nonsense.

We next introduce, following Eddington, the idea, that the universe is a system of N particles, observing Pauli's exclusion principle. We wish to put the energy of the lowest state *of this N-particle system* equal to the *total* rest-energy Nmc^2. It amounts to the same as to order of magnitude (which alone interests us for the moment), if we put mc^2 equal to the energy of the N^{th} (single particle) level. A well known formula

$$N = \frac{4\pi}{3}\frac{V}{\lambda^3} \tag{4}$$

connects the wave length λ of the N^{th} level with N and with the volume $V = 2\,\pi^2 R^3$. From λ follows p by de Broglie's relation : $p = \frac{h}{\lambda} \cdot$

---

(1) « Protons and Electrons », Cambridge 1936.

So we have

$$N = \frac{4\pi}{3}\,\frac{2\pi^2 R^3}{h^3}\,p^3,$$

and therefore

$$\frac{1}{2m}\,p^2 = \frac{1}{2m}\,\frac{h^2 (3N)^{2/3}}{4\,\pi^2 R^2} = mc^2$$

or

$$m = \frac{h\,(3N)^{1/3}}{2\pi Rc\sqrt{2}}\,. \tag{5}$$

The improvement is, that we get no longer R itself, but

$$\frac{R}{N^{1/3}} \approx \frac{h}{mc} \approx 10^{-13}\,\text{cm}\,, \tag{6}$$

that is to say, *the mean interstice* between the particles would have to be of the order of magnitude known to prevail *in the nucleus*. Though this is far from the actual state of the universe (where the average density corresponds to one or two protons per liter), it is rather striking, that we find a great part of the world's material (maybe the bulk of it) in « lumps » of this completely degenerated state. Of course, why the observed mass of a « free » proton, e. g. that of a hydrogen nucleus in CH_4, when investigated in the mass spectrograph, should be essentially determined by the degenerated state in the heavier nuclei, would appear inscrutable.

I should like to mention, that I have tried to arrive at something like formula (5) in many more elaborate and more sophisticated ways than described above, in the faint hope of letting *the square root* of N slip in, instead of the cube root. I have really tried hard. But apart from inappreciable numerical factors I always get the same result.

Eddington arrives at the square root, which makes the formula agree with probable values of R ($\approx 10^{27}$ cm) and N ($\approx 10^{79}$). His main and elaborate derivation, which *accepts* the exclusion principle, is beyond my understanding and I beg to be excused for skipping it. A simple and clear-cut one *foregoes* the exclusion principle, accepts (2), but declares, that the observed rest energy mc^2 corresponds not to the observed particle's zero point energy, but rather to that of the centroid of all the other particles together. (The idea is, I think, that the other particles form the material coordinate frame.) If this is accepted, the desired result follows at once, because, with haphazard directions,

$$p\sqrt{N} = \frac{h}{2\pi R}\sqrt{N}$$

is the momentum of the centroid and now you put, as before,

$$\frac{1}{2m}\left(\frac{h}{2\pi R}\sqrt{N}\right)^2 = mc^2,$$

thus

$$m = \frac{h\sqrt{N}}{2\pi R c\sqrt{2}}.$$

7. *Conclusion.* — These attempts are hardly encouraging. But we must bare in mind the unusual difficulty of the task. In arguing about the Universe, with a view to explain how the laws we know from experiments on a small scale come about, we have nothing to argue with than those small-scale-laws themselves. So we are in the sort of position of that worthy baronet who (said he) dragged himself out of the swamp by pulling his own pigtail.

There is a small length h/mc ($\approx 10^{-13}$ cm), intimately connected with the elementary unit of mass m (that of the proton or neutron) and characteristic of the structure of aggregations of elementary masses (= structure of the nucleus). There is, as a counterpart, a large length $2\pi R$ or R ($\approx 10^{27}$ cm), characteristic of the structure of the Universe. Between the two there is a pretty well established semi-empirical relation, with the number N ($\approx 10^{79}$) of elementary masses that form the Universe intervening, viz.

$$\frac{h}{mc} \approx \frac{R}{\sqrt{N}}. \tag{7}$$

For all that I know, the line of thought reported in the earlier sections of this paper, is the only one to suggest any connection at all between the structure of the Universe as a whole and the structure of the matter that constitutes it. So we have hardly any choice than to follow it, unless we acquiesce in letting sleeping dogs lie.

The Quantization of Gravity

Chapter 17
Without New Difficulties: Quantum Gravity and the Crisis of the Quantum Field Theory Program

Alexander Blum

In the years 1926–1928, immediately following the creation of matrix and wave mechanics, the protagonists of this development elaborated and expanded the techniques of the new quantum mechanics, so as to apply them to field theories.[1] This work culminated in the theory of interacting quantum electrodynamics (QED),[2] published in 1929 by Werner Heisenberg and Wolfgang Pauli (Heisenberg and Pauli 1929a). This paper, which deals mainly with the canonical quantization of both the electromagnetic and the matter-wave fields, famously contains a brief nod to gravitational theory, which is nowadays often quoted jokingly due to its seemingly naive optimism:

> We further note that a quantization of the gravitational field, which appears to be necessary for physical reasons, should be also possible using a formalism entirely equivalent to the one used here without new difficulties. (Heisenberg and Pauli 1929a, 3)

Now this quote doesn't sound half as optimistic if the emphasis is put on *new*. For the theory of quantum electrodynamics which Heisenberg and Pauli had just constructed was replete with difficulties. Three of these difficulties will play an essential role in our story:

1. The theory led to divergent expressions for the energies of stationary states. Even worse, J. Robert Oppenheimer, who was working with Pauli in Zurich at the time, could also show that the *differences* between these energies (i.e., the actually observed frequencies of spectral lines) came out infinite (Oppenheimer 1930).
2. In order to write down a Lorentz-invariant Lagrangian for the interacting electromagnetic and matter-wave fields, it was necessary to work with the electromagnetic potentials ϕ and not just with the fields. But the Lagrangian does not contain then a time derivative of the electric potential ϕ_0, so that there is no corresponding canonical momentum variable, preventing the straightforward implementation of canonical commutation relations.
3. The theory was not manifestly covariant due to the use of equal-time commutation relations. These allowed for a close analogy with the canonical commutation relations of non-relativistic quantum mechanics, but, by singling out time, destroyed

[1] For the early history of quantum field theory, see, e.g., Cini (1982); Darrigol (1982; 1986) and the first two chapters of Schweber (1994). The difficulties of early quantum field theory are discussed in Rueger (1992). An overview that places the development of quantum field theory in the larger context of the development of quantum mechanics can be found in a chapter by Christoph Lehner and the author in a forthcoming volume on the genesis of quantum mechanics.

[2] As opposed to free quantum electrodynamics on the one hand (Jordan and Pauli 1928), and quantum radiation theory, i.e., a theory of transverse photons, on the other (Dirac 1927).

manifest covariance. The Lorentz invariance of the theory thus had to be (and was) proven in a rather roundabout manner.

All three difficulties also played an important role in early work on the quantization of the gravitational field. I will be referring to these three difficulties as the *divergence*, the *momentum*, and the *quantization* difficulty, respectively. I will begin by discussing the *momentum* difficulty.

This difficulty was initially solved by Heisenberg and Pauli by adding to the Lagrangian additional terms, which contained a time derivative of ϕ_0 and were proportional to a parameter ϵ, which was supposed to be set to zero in the final expressions for physical quantities. This procedure was viewed as rather artificial from the start. Heisenberg, who had cooked up the method,[3] described it as a "very crude trick."[4] Heisenberg consequently devised a new method for Heisenberg and Pauli's second paper on QED (Heisenberg and Pauli 1929b).[5] This method relied on the notion of the gauge invariance of the theory of coupled electromagnetic potentials and Dirac matter waves ψ, which Weyl had only recently introduced (see the preceding part), that is, the invariance under a substitution

$$\begin{aligned} \psi(x) &\rightarrow && e^{-\frac{ie}{\hbar c}\chi(x)}\psi(x) \\ \phi_\alpha(x) &\rightarrow && \phi_\alpha(x) + \frac{\partial\chi(x)}{\partial x_\alpha}, \end{aligned} \tag{17.1}$$

where $\chi(x)$ is an arbitrary space-time function. Heisenberg's idea was the following: The field ϕ_0 was simply not quantized, thereby eliminating the need for a canonically conjugate momentum variable in order to construct the canonical commutation relation. ϕ_0 was then simply a (c-number) function of space-time and could be set to zero, due to the well-known underdetermination of the electromagnetic potential. This brought with it a new difficulty, however: With ϕ_0 set to zero, the equation of motion for ϕ_0 (which is simply the first Maxwell equation, or Coulomb's law, $\mathrm{div}\vec{E} = \rho$) no longer resulted from the variation of the Lagrangian and the dynamical problem was underdetermined. The equation of motion could also not simply be added as an operator identity, because it would imply non-vanishing commutation relations between matter and electromagnetic field operators, in contradiction with the canonical commutation relations.

Heisenberg now realized that one had not exploited the full gauge invariance by setting $\phi_0 = 0$. There was still a residual gauge symmetry, since if the function χ doesn't depend on time x_0, the transformation 17.1 leaves ϕ_0 unaltered. To this residual symmetry now corresponded an operator that commutes with the Hamiltonian, that is, a conserved physical quantity. The conserved quantity corresponding to the residual gauge symmetry turned out to be

$$C = \mathrm{div}\vec{E} - \rho. \tag{17.2}$$

One could thus first solve the dynamical problem without the first Maxwell equation and then pick those solutions for which $C = 0$, that is, for which the first Maxwell equation was

[3]See a letter from Pauli to Niels Bohr, 16 January 1929. All the Pauli letters from the 1920s are reprinted in Hermann, Meyenn, and Weisskopf (1979). All translations are by me.

[4]Heisenberg to Jordan, 22 January 1929, Archive for the History of Quantum Physics, MF 18, Section 002–024.

[5]See also a letter from Heisenberg to Pauli, 20 July 1929.

fulfilled at some initial time. There was thus a *Nebenbedingung* (subsidiary condition) on the initial quantum state φ, which had to fulfill the equation $C\varphi = 0$. Since C commutes with the Hamiltonian, this condition on the initial state would propagate, and the first Maxwell equation would always be fulfilled, without actually being an operator identity.

The central difficulty with this new method was the apparent lack of Lorentz invariance: There was now no commutation relation for the zero component of the electromagnetic potential four-vector, and hence the commutation relations were no longer covariant —this was in addition to the difficulty of the equal-time commutators, that is, the *quantization* difficulty. Heisenberg and Pauli convinced themselves that "all statements about gauge invariant quantities [...] fulfill the demand of relativistic invariance," but the proof they presented was highly problematic.[6]

This was the state of affairs for the *momentum* difficulty, when Pauli finally turned to the quantization of gravity. The immediate stimulus appears to have been the reading of Weyl's paper on the interacting theory of Dirac matter waves, electromagnetic potentials and the gravitational field (Chapter 12).[7] He dismissed Weyl's attempt to solve the problem of the negative energy states by using massless Weyl two-spinors instead of Dirac spinors, especially because he did not share Weyl's hope that gravitational effects might be able to generate the mass term:

> The hope of finding a replacement for the mass term in gravitational theory appears illusory to me; the gravitational effects will always be much too small numerically. (Hermann, Meyenn, and Weisskopf 1979, 519)

But Pauli realized that Weyl's general scheme of coupling spinors to a curved space-time metric, by expressing the latter in terms of tetrads, would also work for regular, massive Dirac four-spinors.[8] Reformulated in this manner, Weyl's field theory provided a natural extension of the field theory of interacting matter and electromagnetic waves that Heisenberg and Pauli had quantized for their QED. It thereby provided the obvious starting point for extending and completing their work by also including the gravitational field. Pauli concluded his remarks on Weyl's paper by stating:

> What now interests me most, is the question of how to quantize the $\mathbf{e}(\alpha)$ [the tetrads] themselves in your gravitational theory. (Hermann, Meyenn, and Weisskopf 1979, 520)

This was the question which he set his assistant Léon Rosenfeld. Rosenfeld soon discovered that one encountered problems akin to the *momentum* difficulty of QED.[9] For example, the canonical momenta for the time-components of the four tetrad basis vectors identically vanished. In further pursuing this question, Rosenfeld obtained several interesting results, which he published in a lengthy paper in 1930 (Chapter 18).[10] First of all,

[6] Dirac in a letter to Rosenfeld from 6 May 1932 (Niels Bohr Archive, Copenhagen) (under)stated that he found it "difficult to understand," and Rosenfeld concurred on 10 May (Dirac Papers, Churchill College, Cambridge) that Dirac was "right in not understanding" the "highly doubtful sentence" with which Heisenberg and Pauli concluded their proof.

[7] Letter from Pauli to Weyl, 26 August 1929.

[8] This is of course what Fock had already done. But Pauli was not aware of Fock's work until somewhat later. See a letter to Ehrenfest from 29 September 1929.

[9] Letter from Pauli to Jordan, 30 November 1929.

[10] For a detailed discussion of this work, in particular with regard to its relation with later work on constrained Hamiltonian dynamics, see Salisbury (2009).

he was able to demonstrate in general that the momentum-type difficulties were the result of the invariance of the Lagrangian with regard to certain groups, the gauge group for the case of electrodynamics, the group of general coordinate transformations for general relativity, and the local Lorentz symmetry of the tetrad formulation. He then went on to devise a general method for dealing with such difficulties, a method which also managed to bypass the difficulties of Lorentz covariance encountered in the (second) Heisenberg-Pauli scheme.

In order to sketch Rosenfeld's method, I will focus on the simple case of QED, which is the only example he really worked out to the end. The general idea was to also introduce canonical commutation relations for, and thereby to quantize, the electric potential ϕ_0: Rosenfeld simply assumed that there existed a momentum operator $\mathfrak{P}^4$ that did in fact obey the canonical commutation relation with the electric potential. In order to do this two points needed to be addressed.

First, this meant that the Hamiltonian H would contain a term $\dot{\phi}_0\mathfrak{P}^4$. In order to have the Hamiltonian expressed solely in terms of the canonical variables, one would have to express $\dot{\phi}_0$ in terms of the canonical momenta (and the canonical coordinates, that is, the components of the potential). But the original Lagrangian did not depend on $\dot{\phi}_0$, and consequently $\dot{\phi}_0$ did not show up in the expressions relating the time derivatives of the field and the canonical momenta. This implied that $\dot{\phi}_0$ could be an arbitrary function of space and time without contradicting the defining equations for the canonical momenta. Rosenfeld thus set $\dot{\phi}_0$ in the Hamiltonian equal to an arbitrary function $\lambda(x)$, which then consequently showed up in the equations of motion for the four-potential. In particular the equation of motion for the electric potential was simply of the form $\dot{\phi}_0 = \lambda$, ensuring the self-consistency of the approach. A specification of λ then corresponded to choosing a gauge.

The second point was that one still needed to take account of the fact that the momentum conjugate to the electric potential was actually zero. Rosenfeld introduced a Heisenberg-Pauli type *Nebenbedingung* on the state, demanding that $\mathfrak{P}^4\varphi = 0$. And in order to have this condition propagate in time, an additional condition needed to be imposed

$$0 = \dot{\mathfrak{P}}^4\varphi = \frac{i}{\hbar c}\left[H, \mathfrak{P}_4\right]\varphi = \frac{i}{\hbar c}C\varphi \tag{17.3}$$

One thus obtained the same *Nebenbedingung* that Heisenberg and Pauli had imposed, ensuring the validity of Coulomb's law. No further conditions were necessary, since C was a constant, as Heisenberg and Pauli had already shown.

Since Rosenfeld's scheme was essentially equivalent to the Heisenberg-Pauli method for the specific choice $\lambda = 0$ (although it should be noted that it was not actually necessary to specify λ at all), Rosenfeld's work in the context of QED could simply be viewed as a proof of the covariance of that method, and this is how he later presented it,[11] resorting to the simpler Heisenberg-Pauli method for actual calculations (Rosenfeld 1932).

In any case, it was a whole different approach that came to be the standard method in the QED of the 1930s and 1940s, due to Enrico Fermi, which was based on taking the (Lorenz) gauge condition (as opposed to Coulomb's law) and its time derivative as condi-

[11] Letter to Dirac, 21 May 1932, Dirac Papers, Churchill College, Cambridge.

tions on the wave function (Fermi 1932).[12] It involved the use of a modified Lagrangian that was not gauge invariant and only returned the Maxwell equations if the Lorenz gauge was imposed. Rosenfeld argued against the Fermi approach in his paper, on account of its lacking gauge covariance, demonstrating that his own method was in fact gauge covariant. But this was not a very strong argument at the time, and also Rosenfeld could field no arguments as to why one should attach much weight to gauge covariance in the first place.

How then did Rosenfeld's method fare in the context of gravitation? It did indeed prove applicable, in particular because it was generalizable to the case where the constraints on the canonical momenta were more complicated than the mere vanishing of some of the components, as was the case for the constraints arising from the local Lorentz covariance of the tetrad formalism. But he stopped short of actually constructing a canonical Hamiltonian for the Weyl-Fock field theory. For this, he would have had to solve the equations defining the canonical momenta for the time derivatives of the fields (i.e., of the tetrads), in order to end up with a Hamiltonian that only contained canonical coordinates and momenta, as well as the arbitrary functions of the type λ.

Why did Rosenfeld not do this? This question remains unanswered. Years later, Pauli remarked that Rosenfeld's work was "not satisfactory in all aspects, because he had to introduce certain additional conditions, which no one could really understand."[13] It is entirely unclear, however, what Pauli was referring to. Rosenfeld's work remained for the next two decades an anomaly, until the late 1940s, when the systematic study of constrained Hamiltonian dynamics picked up steam, and several of Rosenfeld's results were re-discovered, as will be discussed in the final part of this book. One additional, simple result of Rosenfeld's, unrelated to the momentum difficulty, has stood the test of time: The necessity of quantizing the gravitational field with commutators, that is, the realization that gravitons must be bosons. Ten years before Pauli's formulation of the spin-statistics theorem, this was a non-trivial result.

One thing is certain: Solving the momentum relations for the time derivatives of the fields is a complicated business. This would not be the last instance of physicists convincing themselves that methods from QED also worked for gravity in principle, without filling in the details. In any case, it turned out to be unnecessary for Rosenfeld's second work on the quantum theory of gravity, which he completed in the same year (Chapter 19). This second paper, much more than the first, which certainly started out as an attempt to quantize gravity, is an attempt to better understand the difficulties of QED by studying whether similar difficulties appeared in other quantum field theories. The difficulty in question here was the first difficulty mentioned above, the *divergence* difficulty, to which we will now turn.

It was initially unclear whether the divergence of the self-energy of the electron in QED was simply an inheritance from the classical theory, where it was well-known that the notion of a point electron was highly problematic, due to the infinite electromagnetic mass associated with such an object. Supposedly Heisenberg had suggested[14] that one

---

[12] It should be noted that Fermi himself actually gave no indications as to how exactly the gauge condition should be interpreted in his original work (Fermi 1929). It was only Heisenberg and Pauli who interpreted it as a condition on the state, akin to their own *Nebenbedingung*, an interpretation which Fermi then adopted.

[13] *... nicht in jeder Hinsicht befriedigend war, da er gewisse zusätzliche Bedingungen einführen musste, die niemand richtig verstehen konnte.* Letter to Oskar Klein, 25 January 1955, reprinted in Meyenn (2001).

[14] When and where is uncertain. Pauli in a letter to Rosenfeld (19 September 1930, but reprinted in Meyenn (1993)) speculates that it might have been during discussions in Copenhagen, but Heisenberg did not actually attend the 1929 spring conference in Copenhagen (8–15 April, see the timeline in Hermann, Meyenn, and

should study the case of the gravitational self-energy of the photon, since there were no singularities in the corresponding classical theory of an electromagnetic wave interacting with the gravitational field. In any case, Pauli egged Rosenfeld on to pursue this problem.[15]

The electromagnetic self-energy of the electron had been calculated perturbatively, expanding in terms of the coupling constant, the electron charge e. The divergent expression then arose at second order in perturbation theory. Consequently, Rosenfeld performed his calculation of the gravitational self-energy of the photon in the approximation of a weak gravitational field. Such an approximation had, of course, been worked out by Einstein immediately after the formulation of general relativity (see the first part of this book). Einstein had started from the assumptions that the deviations $\gamma_{\mu\nu}$ from the flat space-time metric were numerically small (compared to the elements of the flat space-time metric itself, which are of order 1). Rosenfeld now factored out a co-efficient ϵ, the square root of Einstein's gravitational constant, the deviations from the flat space-time metric thus now being of the form $\epsilon\gamma_{\mu\nu}$, Rosenfeld's γ now being a dimensionful quantity. This explicitly made Einstein's approximation equivalent to an expansion in terms of the gravitational coupling constant.

As Einstein had further remarked, solving the field equations of general relativity in this approximation was basically equivalent to calculating retarded potentials in electrodynamics. This also meant that all the techniques developed for QED could be taken over to the quantization of (linearized) gravity in a very straightforward manner. In particular, there was no need for Rosenfeld's new method of dealing with the momentum difficulty.[16] The Fermi method could be taken over from QED and applied directly to gravity, the Lorenz gauge condition being replaced by the analogous coordinate condition imposed by Einstein

$$\frac{\partial \gamma'_{\mu\nu}}{\partial x^\nu} = 0 \tag{17.4}$$

where $\gamma'_{\mu\nu}$ is the trace-reverse of $\gamma_{\mu\nu}$. Following Fermi, Rosenfeld constructed a Lagrangian which was not simply the linearized Einstein-Hilbert Lagrangian, and hence not even approximately invariant under general coordinate transformations. The most straightforward way to arrive at the Rosenfeld Lagrangian is to linearize the first-order Einstein-Hilbert Lagrangian, impose the above coordinate condition and then drop a total divergence. Rosenfeld presented no derivation of his Lagrangian, the only important point being that it returned the correct linearized field equations.

In this framework, Rosenfeld then calculated (to second order in ϵ) the gravitational field energy in the absence of free gravitons, where the operator $\gamma_{\mu\nu}$ can be expressed solely in terms of the electromagnetic field operators. He obtained two divergent terms, one constant and one proportional to the number of photons. The second term was interpreted as the gravitational field energy of a photon, and its divergent nature was subsequently

Weisskopf (1979)), since he was on his world tour (from which he only returned in November, see the timeline in Cassidy (1992)).

[15] Interview of Léon Rosenfeld by Thomas Kuhn and John Heilbron, Niels Bohr Library and Archives, American Institute of Physics, College Park, MD, USA, https://www.aip.org/history-programs/niels-bohr-library/oral-histories/4847-2 (accessed 21 July 2017).

[16] It of course also needs to be noted that there was no need to introduce the tetrads in this case, since no spinorial matter was involved.

cited as proof for the quantum nature of the self-energy divergence, independent of the presence of singularities in the corresponding classical theory.

There remained the suspicion that this divergent energy was merely the coupling of the gravitational field to the infinite zero-point energy of the electromagnetic field. This question was investigated soon after by Jacques Solomon, who redid Rosenfeld's calculation after first eliminating the zero-point energy in the Hamiltonian of the electromagnetic field by imposing normal ordering on the annihilation and creation operators (Solomon 1931). Solomon showed that while the constant divergent term was eliminated in this manner, the photon self-energy remained divergent. But he had only imposed normal ordering on the level of the electromagnetic Hamiltonian and this normal ordering did not survive when calculating the energy of the gravitational field. I therefore think that it can still be argued that Rosenfeld's divergence was merely the result of the coupling to the zero-point energy.

In any case, when Rosenfeld's calculation was redone again by Bryce deWitt (né Seligman) in his unpublished PhD thesis (DeWitt 1949), now using modern covariant techniques, he obtained a vanishing photon self-energy, as should be expected from considerations of gauge invariance alone. DeWitt gave no real argument as to where Rosenfeld had gone wrong, which is quite understandable considering that almost all the calculations from the 1930s were considered obsolete when renormalized quantum field theory was developed in the late 1940s.

The question of the photon self-energy aside, Rosenfeld was still interested in the general question of the quantization of gravity and supplemented his paper with a section on the quantization of the linearized gravitational field also for the case of free gravitons. His treatment was, however, quite sketchy. Bypassing the canonical commutation relations, Rosenfeld went straight to an explicit expression for the operator $\gamma_{\mu\nu}$ in terms of graviton annihilation and creation operators, by expanding in terms of transverse gravitational wave solutions of the free field equation. These transverse waves had been identified by Einstein as the only physical gravitational waves, carrying energy and propagating with the speed of light. Rosenfeld concluded with a brief look at graviton-photon scattering in this linear quantum theory of gravity.

The quantization of linearized GR wasn't worked out in detail until 6 years later, in the PhD work of Soviet physicist Matvei Bronstein (Chapter 20).[17] It is unclear whether he was aware of Rosenfeld's work,[18] but he did not really need to build on it. Bronstein had an independent program of formulating linearized GR as close to electrodynamics as possible, with $\gamma_{\mu\nu}$ treated as a regular field in flat Minkowski space-time (as opposed to as a deviation from the flat space-time metric). He could then straightforwardly apply the canonical, equal-time field quantization procedure developed by Heisenberg and Pauli for QED (along with the Fermi method for dealing with the coordinate condition, which Bronstein now explicitly identified as a gauge condition) to all components of the linearized gravitational field, not only to those that corresponded to freely propagating waves.

These additional components became important when treating the interaction with matter. Here, Bronstein built on Fock's work, expressing Fock's generally relativistic Dirac equation, and in particular the tetrads, in terms of the $\gamma_{\mu\nu}$. Bronstein could thereby not only calculate the emission and absorption coefficients for transverse gravitons, but

[17] On Bronstein's life and scientific work, see Gorelik and Frenkel (1994).

[18] He was certainly aware of Rosenfeld's first paper: During Bronstein's thesis defense (the protocol is reprinted in Kobzarev (1985, 317–320)), Igor Tamm recommended that he make reference to it, which he did in the published Russian paper. Rosenfeld's second paper on linearized gravity is, however, never mentioned.

also show how the non-transverse gravitons led to a Newtonian gravitational interaction between electrons, in the same way that the longitudinal (and time-like) photons of QED led to the Coulomb interaction.

But Bronstein's work is less known for his detailed treatment of the quantum theory of linearized gravity, and more for his musings on the full quantum theory of gravity. Bronstein transferred Bohr and Rosenfeld's analysis of the measurability of field quantities in QED (Bohr and Rosenfeld 1933) to quantum gravity. Bohr and Rosenfeld had concluded that there were no further restrictions on the measurability of the electromagnetic field beside those arising from the canonical uncertainty relations, just as in quantum mechanics. This conclusion had centrally rested on the assumption that the test body used to measure the field was macroscopic (i.e., its atomistic nature could be disregarded) and could be given an arbitrarily high charge density. Bronstein argued that, although the close relation between QED and the quantum theory of gravity allowed adapting almost the entire Bohr-Rosenfeld analysis, the final conclusion would be different, because general relativity did not allow for arbitrarily dense bodies. Rather, there was a fundamental limit from the fact that a body could not become smaller than its gravitational (Schwarzschild) radius. There were thus absolute limits on the measurability of the gravitational field and a quantum theory of gravity would require major conceptual changes.

Now Bronstein's reasoning clearly shows great physical intuition. He was the first to realize the essential difficulties inherent in constructing a quantum theory of gravity and that such a theory was not to be had without fundamental conceptual innovations. His argument is fielded to this day as an elementary demonstration of the problem that is quantum gravity. Why then was Bronstein's argument hardly received by his contemporaries? To understand why it didn't make much of a splash at the time, one needs to look at its weaknesses in somewhat more detail.

One objection was already raised at Bronstein's thesis defense by Wladimir Fock:[19] The absence of a fundamental restriction on the charge density was hardly a desirable feature of quantum electrodynamics. Rather, it allowed for a theory such as QED with point charges that had an apparently infinite charge density, resulting in the divergence difficulties discussed above. Physicists at the time were thus much more inclined to believe that revisions in the concepts of space and time would come from a reformed (quantum) electrodynamics, where a quantity of the order of the classical electron radius would play an essential role. The radical changes that Bronstein was envisioning were thus expected to come in at much larger length scales. It was only when the difficulties of QED (and then also of nuclear physics) had been solved —without radical changes in the underlying space-time theory—that Bronstein's arguments could be viewed as convincing.[20]

But there is a more intrinsic difficulty in Bronstein's argument as well. The aim of the Bohr-Rosenfeld analysis had been to take the established mathematical framework of QED and prove that it was internally consistent, that is, that there were no fundamental limitations on the observability of the electromagnetic field components, which would have been in contradiction with the initial choice of these components as the dynamical (quantum) variables. The Bohr-Rosenfeld analysis was thus a procedure to be applied to an already formulated quantum theory, in order to see whether it contained inherent

[19]The protocol of the defense has been published in Russian, see Footnote 18.

[20]Gorelik and Frenkel are somewhat dismissive of Fock's objection, because they focus too narrowly on a specific attempt at introducing the electron radius into electrodynamics, the Born-Infeld theory, as did Bronstein in his reply to Fock. However, Fock explicitly stated that "the generalzation [of electrodynamics] to a non-linear theory is only beginning" (Kobzarev 1985, 318). He was thus referring to the general program and not merely to its specific, and problematic, implementation by Born and Infeld.

contradictions. Consequently, in the original Russian article based on his thesis (Bronstein 1936), Bronstein applied this analysis to his linearized quantum gravity (which was a fully formulated quantum field theory), and not to an as yet hypothetical full quantum theory of gravity. He then voiced the hope that his argument would also carry over to the full theory, even though his analysis could not be applied to it, since it had not yet been formulated.

Somewhere between writing up his thesis and publishing the German-language article reprinted here, Bronstein must have realized that this argument was fundamentally flawed. There was nothing in the classical, linearized theory that prevented a body from being smaller than its gravitational radius. And so the linearized quantum theory was just as consistent as QED. In the German article, he thus presented the Bohr-Rosenfeld argument as an argument against the coherence of a quantum field theory based on the full non-linear theory of gravity, disregarding the obvious logical difficulty of showing the inconsistency of a theory that didn't even exist yet.

This criticism of Bronstein's argument was raised by Jacques Solomon (Chapter 21), but not without bringing forth a further argument against the possibility of constructing a quantum field theory based on the full non-linear theory.[21] Solomon's argument was based on a recent proof by Nathan Rosen on the non-existence of non-singular plane wave solutions in full GR (Rosen 1937).[22] Now, all quantum field theories of the time were based on an expansion of the field quantities in plane waves. Even if Heisenberg and Pauli's starting point had been commutation relations for the field quantities themselves, as soon as they went to the level of representing these field quantities as operators on a Hilbert space, they had, as they grudgingly admitted, no choice but to expand them in plane wave solutions of the free field equations and treat the expansion coefficients as annihilation and creation operators on occupation number space. From the absence of plane wave solutions, Solomon concluded that the *present* formalism of field quantization was not compatible with the non-linear theory of gravitation, a weaker claim than Bronstein's, but one that found at least one interested listener in Pauli, as we shall see later.

But of course both Bronstein's and Solomon's thoughts on the difficulties of the non-linear quantum theory of gravity could be regarded as nitpicking, as long as the difficulties of QED (and thereby also of the quantum theory of linearized gravity) remained unsolved. The attempts to solve these difficulties, in as far as they were difficulties of quantum field theory in general, of course also impacted the pursuit of quantum gravity. It is well-known that the most fundamental difficulty of QED, the *divergence* difficulty, was not solved until the late 1940s. But progress was being made already in the 1930s on lesser difficulties, in particular on the lack of manifest covariance due to the use of equal time commutators, what I have called the *quantization* difficulty. In the last part of this essay, I will focus on this third and final difficulty.

While certainly not the most pressing difficulty at the time, there were enough physicists who believed that formulating quantum theory in a more overtly relativistic manner was a worthwhile endeavor. Two important relativistic quantization procedures were devised in order to replace equal-time commutators in the first half of the twentieth century.[23] In both cases, Paul Dirac played an essential role. The union of relativity and quantum theory was a leitmotif in his work from the very start, when he attempted to make Heisen-

---

[21] In the secondary literature, e.g. in Stachel (1999), Solomon's and Bronstein's arguments have in general been lumped together.

[22] Rosen's conditions were later seen as too strict, since he was also ruling out mere coordinate singularities (Bondi, Pirani, and Robinson 1959).

[23] Not counting Feynman's path integral, which was not applied to the case of gravitation until the mid-1950s, see in particular Misner (1957).

berg's matrix mechanics relativistic by turning time into a non-commuting matrix (Dirac 1926). This was followed by the Dirac equation in 1928 (Dirac 1928) and then by several hugely influential papers in 1932/33, in which he laid the foundations for several new, relativistically invariant quantization techniques. We begin by discussing his 1933 paper on "The Lagrangian in Quantum Mechanics" (Dirac 1933).

Dirac's general idea was to formulate quantum mechanics not in terms of states (wave functions) propagating in time, but in terms of transition amplitudes from the state at one time to the state at a later time. Dirac called these transition amplitudes "transformation functions," since he conceptualized them as generating canonical transformations from the canonical coordinates at one time to those at a later time. In the final section, he also hinted at how to generalize this idea to quantum field theory, where the initial and final times would be replaced by an arbitrary (not necessarily space-like) three-dimensional hypersurface of space-time, which formed the integration boundary of the classical action. He dubbed the field theoretic amplitudes relating the canonical (field) coordinate values on different points on this hypersurface[24]"generalized transformation functions." Dirac's ideas, however, remained very vague and were not worked out in any detail by him.

Dirac's idea was taken up by Paul Weiss,[25] who realized that the procedure of replacing initial and final times by general three-dimensional boundary hypersurfaces might also be applicable to the canonical quantization procedure (Weiss 1936). At first sight, the canonical quantization procedure seemed to rest essentially on the singling out of time, since the momentum canonically conjugate to a field variable ϕ appearing in the Lagrangian L was defined as $\frac{\partial \mathscr{L}}{\partial(\partial_t \phi)}$. This was one of the reasons why Dirac had based his sketch on the Lagrangian formalism, where the canonical momenta do not enter. Weiss now noted that the canonical momentum could also be defined in another way. If one reads the definition of the canonical momentum as the time component of the four-vector $\frac{\partial \mathscr{L}}{\partial(\partial_\mu \phi)}$, then this can be generalized by taking the canonical momentum as the component of the four-vector orthogonal to the hypersurface forming the integration boundary of the action, on which the values of the field variables are fixed.

Weiss initially, like Dirac, took the boundary hypersurface to be closed. So, while one could locally determine the direction w perpendicular to the surface and describe points on the surface by three parameters u_i, this coordinate system (dubbed "natural" by Weiss) could not be regular throughout space-time. Still, one could write down commutators for the field operators, which were exactly of the form of the Heisenberg-Pauli equal-time commutators, except that the arguments of the field operators were now not the three spatial coordinates, but rather the three surface parameters u_i. Further, a Hamiltonian could be constructed in the usual fashion, using the generalized canonical momenta, and be shown to generate the "time evolution" (i.e., the evolution for increasing w) of the field operators.

In his second, and most influential paper (Chapter 22) , Weiss could show that the generalization of the commutation relations need not necessarily occur only at the quantum level, where it smacked of an ad-hoc modification of the well-established canonical quantization procedure. Instead, he could demonstrate that his generalized commutation relations corresponded already at the classical level to generalized Poisson brackets, and that his method was thus a straightforward generalization of the usual canonical quantization method.

[24]They should not be called transition amplitudes, since this would imply a notion of initial and final state. Maybe "correlation amplitudes" would be a good name.
[25]For a detailed discussion of Weiss's life and work, see Rickles and Blum (2015).

This reconceptualization brought with it a slight modification of the original method. If the generalized Poisson brackets are supposed to be invariant under canonical transformations, in particular under time evolution, as are their point-mechanical counterparts, they can only sensibly be defined on some initial space-like hypersurface (on which then the values of both the field variables and the canonical momenta are given). Weiss thus dropped his original approach of having the commutation relations defined on an entire closed hypersurface, and replaced it with commutation relations on some (initial) space-like hypersurface, which corresponded to the classical, generalized Poisson bracket. This quantization procedure provided a covariant generalization of the equal-time commutators that, however, still stuck closely to the canonical scheme. We will discuss its further use in the last part of this book, which discusses the beginnings of what would come to be known as the canonical quantization approach to quantum gravity.

The other quantization procedure that went beyond the Heisenberg-Pauli equal-time commutators was not a direct generalization of the canonical quantization procedure, as Weiss's method was. At its core were not the commutation relations between canonically conjugate field variables, but rather the commutation relations between two field variables at arbitrary different points in space-time. Such covariant commutators had first been used in 1928 by Jordan and Pauli for their quantization of the free electromagnetic field (Jordan and Pauli 1928). They were constructed by expanding the electromagnetic field variables in terms not of time-independent spatial modes, but in terms of time-dependent, propagating plane waves, which were solutions of the (free) equations of motion. Imposing the usual annihilation-creation operator commutation relations on the expansion coefficients, and then re-summing, the covariant commutation relations were obtained.

Since the construction was based on an expansion in terms of solutions of the wave equation, the commutation relations were automatically compatible with the equations of motion, a necessary requirement for any equation relating the field at two time-like (or, for the case of electrodynamics more importantly, light-like) separated points. Indeed, the commutators themselves actually had to be (singular) solutions of the equations of motion. This immediately became intractable once one was dealing with interacting fields and non-linear equations of motion, which is why one year later, Pauli reverted to the use of equal-time commutators.

Covariant commutation relations were brought back into the game by Dirac in 1932. In an attempt to relaunch QED, he had effectively reconstructed Heisenberg-Pauli quantum field theory, only now in the interaction representation (Dirac 1932). Dirac was much ridiculed by Pauli for having merely produced an equivalent theory, despite his grandiloquent claims to the contrary. In a letter to Dirac of 11 September 1932,[26] Pauli wrote:

> Your remarks on quantum electrodynamics [...] were—to put it mildly—certainly no masterpiece. After a confused introduction, consisting of sentences that were only halfway understandable because they were only halfway understood, you finally arrive at results for a simplified, one-dimensional example that are identical with those obtained by applying the formalism of Heisenberg and myself to this example. (The identity is immediately recognizable and was then calculated in too complicated a manner by Rosenfeld.) This conclusion of your work stands in contradiction to the more or less clearly voiced claims in the introduction that you would somehow be able to make a better quantum electrodynamics than Heisenberg and myself. (Hermann, Meyenn, and Weisskopf 1985, 115)

[26] All Pauli letters from the 1930s are reprinted in Hermann, Meyenn, and Weisskopf (1985).

But when Dirac, together with Vladimir Fock and Boris Podolsky, actually applied his interaction representation formulation of quantum field theory to the full electrodynamics (Dirac, Fock, and Podolsky 1932), and not just to the toy example of the first paper, Pauli realized that it had a huge advantage: All the non-trivial dynamics due to the interaction between charged matter and electromagnetic field were relegated to the time evolution of the state vector. In other words: Also the second inhomogeneous Maxwell equation, the Ampère-Maxwell law, was satisfied only through the action of the field operators on the wave function and not as an operator identity. This implied that the electromagnetic field operators still obeyed the free field equations and consequently the covariant commutation relations of Jordan and Pauli. On 2 June 1933, Pauli wrote to Heisenberg:

> As time goes by, I find myself liking the work of Dirac, Fock and Podolsky more and more. It is funny that there the commutation relations of vacuum electrodynamics do not change when particles are present, even for $t \neq t'$. (Hermann, Meyenn, and Weisskopf 1985, 167)

But the method of quantizing by imposing covariant commutation relations was still an isolated technique, applied only to the Maxwell field, far-removed from the generality of the canonical quantization approach, which could be applied to any classical field theory given in Lagrangian or Hamiltonian form. It was extended in 1938 by Stueckelberg to encompass massive scalar fields, again by explicitly expanding the field operators in terms of solutions of the wave equation (in this case the Klein Gordon equation) (Stückelberg 1938a), and then also to massive vector (Proca) fields (Stückelberg 1938b). In 1939, finally, Markus Fierz gave a general expression for the covariant commutation relations for fields with arbitrary spin (Fierz 1939).[27] The covariant quantization method thereby became a full-fledged quantization method in its own right, applicable to any Lorentz-covariant field theory.

What did this mean for the quantization of gravity? General Relativity was, of course, not a Lorentz-invariant field theory, but as we have seen, the linearized theory could be thought of in this manner. And indeed, Fierz together with Pauli could show that the linearized gravitational field fit into Fierz's scheme, namely as a massless spin 2 field (Chapter 23).

This was a major conceptual change: Bronstein (and Rosenfeld) had started from the full theory of general relativity and obtained a Lorentz-covariant field theory as an approximation. Pauli and Fierz now showed that one could construct the field theory of linearized gravity equally well *without* any reference to general relativity: it was the unique Lorentz-covariant field theory of a massless spin 2 field. Also the transformation properties of the metric perturbations $\gamma_{\mu\nu}$ arose, without any reference to general coordinate transformations, as the gauge transformation properties of a massless field, a straightforward generalization of electrodynamics.

Initially, of course, this change was hardly consequential: The quantum field theory of linearized gravity had already been worked out, and neither the use of covariant quantization methods, nor the reconceptualization in terms of a spin 2 field made any difference to, say, Bronstein's formulation. But it would eventually lead to a new way of thinking about the quantum theory of non-linear gravity. Pauli, when recapitulating his work with Fierz in his manuscript for the 1939 Solvay conference (which never took place due to the outbreak of World War II, but is reprinted in Meyenn (1993)), was still quite pessimistic

[27] The context of these works is the discovery of the putative Yukawa meson in 1937, see Blum (2014).

concerning the prospects for the non-linear theory. In particular he cited Solomon's argument, but with an interesting twist:

> It is certainly a limitation of the quantum theoretical side of these considerations that one here contents oneself with the approximation in which the generally relativistic field equations are linear. This limitation is intimately connected with the well-known divergence difficulties of field theory. (Meyenn 1993, 901)

No further explanation of this connection is given, but I would suggest the following reading: In quantum electrodynamics it is the non-linear terms, that is, the terms coupling the electromagnetic field to the charged current of the electrons, that lead to the divergence difficulties. Pauli's connection of the non-linearities of GR with the divergence difficulties of QED thus offered an entirely new way to look at these non-linearities: For Solomon, these were terms arising in a higher approximation to the full theory, spoiling the usual quantum field theoretical approach by eliminating the possibility of plane wave solutions. In contrast, the non-linear terms could now be considered as interaction terms, added on to the Lagrangian for the free spin 2 field, just like the interaction terms of QED were added on to the free field Lagrangians, with the sole difference that one was now talking about self-interactions of a single field.

As long as the divergence difficulties of QED remained unresolved, this hardly made a difference. But when covariant renormalization was developed in the late 1940s, it seemed plausible that the same techniques might also be used for the non-linear theory. By using the interaction representation, one could quantize the gravitational field in the same manner as the electromagnetic field, by separating the free spin 2 quantum (which had plane wave solutions, thereby circumventing Solomon's objection) from its self-interaction. The self-interaction (as well as the interactions with the matter fields) could then be treated in the same manner as the interactions of QED, possible divergences being absorbed into the properties of the free field. Thus the program of covariant quantization was born, which, however, took off only somewhat after the period covered in this book. We thus see how the two major programs for the quantization of gravity, canonical and covariant quantization,[28] had their origins in the 1930s, in two attempts to get over the *quantization* difficulty of QED with, at this early stage, only weak and incidental connections to the specific problem of quantum gravity.

References

Blum, Alexander (2014). From the Necessary to the Possible: The Genesis of the Spin-Statistics Theorem. *European Physical Journal H* 39(5):543–574.

Bohr, Niels and Leon Rosenfeld (1933). Zur Frage der Messbarkeit der elektromagnetischen Feldgrößen. *Det Kgl. Danske Videnskabernes Selskab Mathematisk-fysiske Meddelelser* 12:3–65.

Bondi, Hermann, Felix A. E. Pirani, and Ivor Robinson (1959). Gravitational Waves in General Relativity. III. Exact Plane Waves. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 251(1267):519–533.

Bronstein, Matwei Petrowitsch (1936). Kvantovanie Gravitatsionn'ikh Woli. *Zhurnal Eksperimental'noj i Teoreticeskoj Fiziki* 6(3):195–236.

[28] See, e.g.,Rovelli (2004), Appendix B.

Cassidy, David (1992). *Uncertainty: The Life and Science of Werner Heisenberg*. New York: W. H. Freeman.

Cini, Marcello (1982). Cultural Traditions and Enviromental Factors in the Development of Quantum Electrodynamics (1925–1933). *Fundamenta Scientiae* 3:229–253.

Darrigol, Olivier (1982). *Les débuts de la théorie quantique des champs. 1925–1948*. PhD thesis. Université de Paris I.

— (1986). The Origin of Quantized Matter Waves. *Historical Studies in the Physical and Biological Sciences* 16:197–253.

DeWitt, Bryce (Dec. 1949). *I: The Theory of Gravitational Interactions. II: The Interaction of Gravitation with Light*. Published as Carl Bryce Seligman. PhD thesis. Harvard.

Dirac, Paul A. M. (1926). Relativity Quantum Mechanics with an Application to Compton Scattering. *Proceedings of the Royal Society, Series A* 111(758):405–423.

— (1927). The Quantum Theory of Emission and Absorption of Radiation. *Proceedings of the Royal Society of London* A114:243–265.

— (1928). The Quantum Theory of the Electron. *Proceedings of the Royal Society of London: Series A, Containing Papers of a Mathematical and Physical Character*: 610–624.

— (1932). Relativistic Quantum Mechanics. *Proceedings of the Royal Society, Series A* 136(829):453–464.

— (1933). The Lagrangian in Quantum Mechanics. *Physikalische Zeitschrift der Sowjetunion* 3(1):64–72.

Dirac, Paul A. M., Vladimir A. Fock, and Boris Podolsky (1932). On Quantum Electrodynamics. *Physikalische Zeitschrift der Sowjetunion* 2(6):468–479.

Fermi, Enrico (1929). Sopra l'elettrodinamica quantistica. *Rendiconti Lincei* 9:881–887.

— (1932). Quantum Theory of Radiation. *Review of Modern Physics* 4:87–132.

Fierz, Markus (1939). Über die relativistische Theorie kräftefreier Teilchen mit beliebigem Spin. *Helvetica Physica Acta* 12:3–37.

Gorelik, Gennady E. and Victor Ya. Frenkel (1994). *Matvei Petrovich Bronstein and Soviet Theoretical Physics in the Thirties*. Basel: Birkhäuser.

Heisenberg, Werner and Wolfgang Pauli (1929a). Zur Quantendynamik der Wellenfelder. *Zeitschrift für Physik* 56:1–61.

— (1929b). Zur Quantentheorie der Wellenfelder. Band II. *Zeitschrift für Physik* 59(3–4):168–190.

Hermann, Armin, Karl von Meyenn, and Victor F. Weisskopf, eds. (1979). *Wolfgang Pauli: Wissenschaftlicher Briefwechsel mit Bohr, Einstein, Heisenberg u.a.* Vol. 1: 1919–1929. New York: Springer.

— eds. (1985). *Wolfgang Pauli: Wissenschaftlicher Briefwechsel mit Bohr, Einstein, Heisenberg u.a.* Vol. 2: 1930–1939. New York: Springer.

Jordan, Pascual and Wolfgang Pauli (1928). Zur Quantenelektrodynamik ladungsfreier Felder. *Zeitschrift für Physik* 47:151–173.

Kobzarev, Igor′ Ju., ed. (1985). *Einsteinovski Sbornik, 1980–1981*. Moscow: Nauka.

Meyenn, Karl von, ed. (1993). *Wolfgang Pauli: Wissenschaftlicher Briefwechsel mit Bohr, Einstein, Heisenberg u.a.* Vol. 3: 1940–1949. Berlin: Springer.

— ed. (2001). *Wolfgang Pauli: Wissenschaftlicher Briefwechsel mit Bohr, Einstein, Heisenberg u.a.* Vol. 4, Part 3: 1955–1956. Berlin: Springer.

Misner, Charles W. (1957). Feynman Quantization of General Relativity. *Reviews of Modern Physics* 29(3):497–509.

Oppenheimer, J. Robert (1930). Note on the Theory of the Interaction of Field and Matter. *Physical Review* 35(5):461–477.

Rickles, Dean and Alexander Blum (2015). Paul Weiss and the Genesis of Canonical Quantization. *European Physical Journal H* 40(4–5):469–487.

Rosen, Nathan (1937). Plane Polarised Waves in the General Theory of Relativity. *Physikalische Zeitschrift der Sowjetunion* 12:366–372.

Rosenfeld, Leon (1932). La Théorie quantique des champs. *Annales de l'Institut Henri Poincaré* 2(1):25–91.

Rovelli, Carlo (2004). *Quantum Gravity*. Cambridge: Cambridge University Press.

Rueger, Alexander (1992). Attitudes towards Infinities: Responses to Anomalies in Quantum Electrodynamics, 1927–1947. *Historical Studies in the Physical and Biological Sciences* 22:309–337.

Salisbury, Donald (2009). Léon Rosenfeld and the Challenge of the Vanishing Momentum in Quantum Electrodynamics. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 40(4):363–373.

Schweber, Silvan S. (1994). *QED and the Men Who Made It: Dyson, Feynman, Schwinger, and Tomonaga*. Princeton: Princeton University Press.

Solomon, Jacques (1931). Nullpunktsenergie der Strahlung und Quantentheorie der Gravitation. *Zeitschrift für Physik* 71(3–4):162–170.

Stachel, John (1999). The Early History of Quantum Gravity (1916–1940). In: *Black Holes, Gravitational Radiation and the Universe*. Springer, 525–534.

Stückelberg, Ernst C. G. (1938a). Die Wechselwirkungskräfte in der Elektrodynamik und in der Feldtheorie der Kernkräfte: Teil 1. *Helvetica Physica Acta* 11(225–244).

— (1938b). Die Wechselwirkungskräfte in der Elektrodynamik und in der Feldtheorie der Kernkräfte: Teil 1fte in der Elektrodynamik und in der Feldtheorie der Kernkräfte: Teil 2 und 3. *Helvetica Physica Acta* 11:299–328.

Weiss, Paul (1936). On the Quantization of a Theory Arising from a Variational Principle for Multiple Integrals with Application to Born's Electrodynamics. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 156(887): 192–220.

Chapter 18
Léon Rosenfeld (1930): Zur Quantelung der Wellenfelder

Léon Rosenfeld (1930). Zur Quantelung der Wellenfelder. *Annalen der Physik*, 397: 113–152.

Zur Quantelung der Wellenfelder

Von L. Rosenfeld

Einleitung

Wesentliche Fortschritte in der Formulierung der allgemeinen Quantengesetze der elektromagnetischen und materiellen Wellenfelder haben neuerdings Heisenberg und Pauli[1]) erzielt, indem sie die von Dirac erfundene „Methode der nochmaligen Quantelung“ systematisch entwickelten. Neben gewissen sachlichen Schwierigkeiten, die viel tiefer liegen, trat dabei eine eigentümliche Schwierigkeit formaler Natur auf: der zum skalaren Potential kanonisch konjugierte Impuls verschwindet identisch, so daß die Aufstellung der Hamiltonschen Funktion und der Vertauschungsrelationen nicht ohne weiteres gelingt. Zur Beseitigung dieser Schwierigkeit sind bisher drei Methoden vorgeschlagen worden, die zwar ihren Zweck erfüllen, aber doch schwerlich als befriedigend betrachtet werden können.

1. Die erste Heisenberg-Paulische Methode ist ein rein analytischer Kunstgriff.[2]) Man fügt zur Lagrangefunktion gewisse Zusatzglieder hinzu, die mit einem kleinen Parameter ε multipliziert sind und bewirken, daß der obenerwähnte Impuls nicht mehr verschwindet. In den Schlußresultaten muß man dann zum Limes $\varepsilon = 0$ übergehen. Die ε-Glieder führen aber zu unphysikalischen Rechenkomplikationen[3]) und zerstören die charakteristische Invarianz der Lagrangefunktion gegenüber der Eichinvarianzgruppe.

2. Die zweite Heisenberg-Paulische Methode[4]) benutzt hingegen wesentlich diese Invarianz. Dem skalaren Potential

1) W. Heisenberg u. W. Pauli, Ztschr. f. Phys. **56**. S. 1. 1929; ebenda **59**. S. 168. 1930. Im folgenden mit H. P. I bzw. II zitiert.

2) H. P. I, S. 24—26, 30ff.

3) Vgl. L. Rosenfeld, Ztschr. f. Phys. **58**. S. 540. 1929.

4) H. P. II.

wird ein bestimmter, beliebiger Wert, z. B. Null, gegeben; dann liefert die Hamiltonsche Methode eine Bewegungsgleichung weniger. Lautet nun die fehlende Gleichung $C = 0$, so findet man auf Grund der Eichinvarianz der Hamiltonfunktion $C =$ konst. Die Wahl des Wertes 0 für diese Konstante bedeutet die Beschränkung auf eines von verschiedenen untereinander nicht kombinierenden Termsystemen. Das Auszeichnen einer Komponente des Viererpotentials bringt aber mit sich die Notwendigkeit eines Beweises für die relativistische Kovarianz des Verfahrens; und dieser Nachweis ist sehr mühsam.

3. Die Fermische Methode[1]) besteht auch im Hinzufügen von Zusatzgliedern zur Lagrangefunktion derart, daß kein Impuls mehr identisch verschwindet. Damit die so erhaltenen Feldgleichungen mit den gewöhnlichen übereinstimmen, müssen gewisse Nebenbedingungen erfüllt sein; es muß dann gezeigt werden, daß, wenn diese Nebenbedingungen auf einem Schnitt $t =$ const gelten, sie sich dann von selbst im Laufe der Zeit fortpflanzen. Der Nachteil dieser Methode ist der, daß wiederum die Eichinvarianz zerstört wird.

Nun ist das identische Verschwinden der genannten Impulskomponente keineswegs eine vereinzelte Erscheinung; denn der Grund dafür ist eben die Eichinvarianz der Lagrangefunktion, wie eine leichte, weiter unten ausführlich dargelegte Überlegung zeigt. Analoges, d. h. allgemeiner das Auftreten von identischen Relationen zwischen den Variablen und den konjugierten Impulsen, ist in allen Fällen zu erwarten, wenn die Lagrangefunktion eine geeignet gebaute Gruppe gestattet. Bei der näheren Untersuchung dieser Verhältnisse an Hand des besonders lehrreichen Beispieles der Gravitationstheorie, wurde ich nun von Prof. Pauli auf das Prinzip einer neuen Methode freundlichst hingewiesen, die es in durchaus einfacher und natürlicher Weise gestattet, das Hamiltonsche Verfahren beim Vorhandensein von Identitäten auszubilden, ohne den Nachteilen der bisherigen Methoden ausgesetzt zu sein. Im folgenden wird der Gegenstand zunächst vom all-

---

1) Vgl. H. P. II, S. 171, Fußnote.

gemeinen gruppentheoretischen Standpunkt behandelt, sodann an Hand verschiedener physikalischer Beispiele illustriert.[1])

Erster Teil: **Allgemeine Theorie**

§ 1. Ansätze über die Lagrange-Funktion und die zugrunde gelegte Gruppe

Wir betrachten irgendein dynamisches System, definiert durch Feldgrößen Q_α (x^1, x^2, x^3, x^4), welche von den Raumkoordinaten x^1, x^2, x^3 und von der Zeitkoordinate $x^4 = ct$ (und *nicht*, wie bei H. P., $x^4 = i\,c\,t$!) abhängen. Über die Lagrangefunktion $\mathfrak{L}$ $(Q;\ \partial Q / \partial x)$ brauchen wir keine Annahme zu machen, solange wir im Rahmen der klassischen Theorie bleiben, d. h. mit lauter c-Zahlen operieren; betrachten wir aber die Variabeln Q_α als q-Zahlen (während die Raumzeitkoordinaten immer c-Zahlen bleiben), so müssen wir berücksichtigen, daß der Satz von der Ableitung einer Funktionenfunktion seine allgemeine Gültigkeit verliert[2]); wollen wir also (und das wird der Fall sein) gewisse Eigenschaften der Lagrangefunktion, die aus diesem Satz fließen, beibehalten, so sind wir genötigt, über die Funktion $\mathfrak{L}$ solche einschränkenden Annahmen zu machen, daß die betreffenden Eigenschaften trotz des Versagens des genannten Satzes gültig bleiben. Es zeigt sich nun, daß diese Einschränkungen zwar vom mathematischen Standpunkt sehr weitgehend sein müssen, daß sie jedoch bei den physikalisch interessanten Lagrangefunktionen erfüllt sind. Sie betreffen einmal die analytische Beschaffenheit der Lagrangefunktion: diese soll höchstens quadratisch in den Ableitungen der Q_α sein; ferner die Reihenfolge der miteinander nicht vertauschbaren Größen.

Zur Abkürzung schreiben wir oft $Q_{\alpha,\nu}$ statt $\dfrac{\partial Q_\alpha}{\partial x^\nu}$, auch $\dot{Q}_\alpha$ statt $Q_{\alpha,4} \equiv \dfrac{\partial Q_\alpha}{\partial x^4}$. Ferner unterdrücken wir die Summa-

1) Hier möchte ich ein für allemal betonen, daß die in den Arbeiten H. P. I und II behandelten Spezialfälle mir oft den Weg zur gewünschten Verallgemeinerung zeigten. Es hätte wenig Zweck, im folgenden jedesmal darauf hinzuweisen.

2) Vgl. H. P. I, S. 18, ferner S. 14, Fußnote 1.

tionszeichen gemäß der bekannten Regel. Mit diesen Festsetzungen lautet nun unser Ansatz für die Lagrangefunktion:

$$(1)\quad 2\mathfrak{L} = Q_{\alpha,\nu}\,\mathfrak{A}^{\alpha\nu,\beta\mu}(Q)\,Q_{\beta,\mu} + Q_{\alpha,\nu}\,\mathfrak{B}^{\alpha\nu}(Q) + \mathfrak{B}^{\alpha\nu}(Q)\,Q_{\alpha,\nu} + \mathfrak{C}(Q).$$

Obwohl nur die $\dot{Q}_\alpha$ mit den Q_α nicht vertauschbar sind, müssen wir doch auch für die anderen Ableitungen an einer bestimmten Reihenfolge festhalten; denn gewisse Operationen, z. B. d/dx^4, verwandeln die betreffenden Größen in andere, die untereinander nicht mehr vertauschbar sind, so daß das Resultat einer solchen Operation von der ursprünglichen Reihenfolge abhängt.

Da die c-Zahlüberlegungen oft an Allgemeinheit und Eleganz überlegen sind, wollen wir sie im folgenden zuweilen zur ersten Übersicht benutzen und nachher die für die q-Zahltheorie erforderlichen Modifikationen andeuten. Um aber unnötige Wiederholungen zu vermeiden, reden wir auch bei c-Zahlen von Vertauschungsrelationen, wobei wir natürlich die korrespondierenden Poissonschen Klammersymbole meinen.

Nun zur Definition der Transformationsgruppe, welche die Lagrangefunktion (in einem näher zu präzisierenden Sinne) gestatten soll. Es liegt uns in dieser Untersuchung keineswegs daran, die größtmögliche Allgemeinheit anzustreben, sondern die Darstellung nur so allgemein zu halten, daß in den physikalischen Anwendungen die tieferen Zusammenhänge klar hervortreten. Wir fragen also nicht nach der allgemeinsten Gruppe, die bei gegebener Lagrangefunktion Identitäten von der oben besprochenen Art zur Folge hat, sondern wir legen eine speziellere, wenn auch ausgedehnte, Klasse von kontinuierlichen unendlichen Gruppen zugrunde, von der wir zeigen, daß sie bei beliebiger Lagrangeschen (c-Zahl-)Funktion zu Identitäten führen.[1])

Wir charakterisieren unsere Gruppe durch ihre infinitesimale Transformation; wir nehmen an, daß sich sowohl die x^ν wie die Q_α auf bestimmte Weise transformieren, und zwar

1) Die dabei benutzte Methode gibt übrigens sofort die Antwort auf die eben aufgeworfene allgemeine Frage. Bei speziellem Bau der Lagrangefunktion braucht die Gruppe sogar nicht unendlich zu sein, um Identitäten zu bedingen.

hängen die δx^ν bzw. δQ_α ab von r_0 willkürlichen reellen Funktionen $\xi^r(x)$ $(r = 1, 2, \ldots r_0)$ und ihren Ableitungen bis zur Ordnung k bzw. j; die Koeffizienten dieser Ableitungen sollen reell sein und (hierin liegt die Spezialisierung der Gruppen) in δx^ν nur von den x^ν, in δQ_α nur von den x^ν und den Q_α (und nicht von den Ableitungen der Q_α) abhängen. In Formeln:

$$(2)\quad \begin{cases} \delta x^\nu = a_r^{\nu,0}(x)\,\xi^r(x) + a_r^{\nu,\sigma}(x)\dfrac{\partial \xi^r}{\partial x^\sigma} + a_r^{\nu,\sigma\ldots\tau}(x)\dfrac{\partial^k \xi^r}{\partial x^\alpha \cdots \partial x^\tau}, \\ \delta Q_\alpha = c_{\alpha r}^0(x, Q)\,\xi^r(x) + c_{\alpha r}^\sigma(x, Q)\dfrac{\partial \xi^r}{\partial x^\sigma} \\ \qquad\qquad + c_{\alpha r}^{\sigma\ldots\tau}(x, Q)\dfrac{\partial^j \xi^r}{\partial x^\sigma \cdots \partial x^\tau}. \end{cases}$$

Dazu kommt noch die wesentliche Voraussetzung, daß[1])

$$(3)\qquad j \geqslant k + 1.$$

Was die Vertauschungseigenschaften der in (2) vorkommenden Funktionen betrifft, so sollen die ξ^r c-Zahlen sein und diese Eigenschaft bei allen Transformationen der Gruppe (2) behalten (wie es ja der Festsetzung, daß die ξ^r nur von den x^ν abhängen, entspricht). Da die a nur von den x^ν abhängen, dürfen wir sie gleichfalls als c-Zahlen betrachten. Dann sind auch die δx^ν c-Zahlen, wie es sein muß, damit wir die x^ν selber als c-Zahlen behandeln dürfen.

Die wichtigsten in der Physik vorkommenden Gruppen fallen unter diesen Typus (vgl. den zweiten Teil dieser Arbeit).

Es bleibt jetzt noch übrig, auszudrücken, daß das Integral

$$\int \mathfrak{L}\, dx^1\, dx^2\, dx^3\, dx^4$$

bei den Transformationen (2) invariant bleibt. Zu dem Zwecke führen wir zunächst einige Begriffe ein.

Neben die „lokale" Variation $\delta \Phi\,(x, Q, \partial Q/\partial x, \ldots)$ tritt die „substantielle" Variation

$$(4)\qquad \delta^* \Phi = \delta \Phi - \frac{d\Phi}{dx^\nu}\,\delta x^\nu;$$

---

1) Wir setzen $\dfrac{\partial^0 \xi}{(\partial x)^0} \equiv \xi$ und $\dfrac{\partial^{-1} \xi}{(\partial x)^{-1}} \equiv 0$.

wenn wir die transformierten Größen mit einem Strich versehen, so ist

$$\delta \Phi = \Phi'[x'; Q'(x'); \ldots] - \Phi[x; Q(x); \ldots],$$

während

$$\delta^* \Phi = \Phi'[x; Q'(x); \ldots] - \Phi[x; Q(x); \ldots]$$

bedeutet. Daraus folgen unmittelbar die wichtigen, auch für q-Zahlen gültigen Formeln:

$$\delta^* \frac{d\Phi}{dx^\nu} = \frac{d}{dx^\nu} \delta^* \Phi, \tag{5}$$

$$\delta \frac{d\Phi}{dx^\nu} = \frac{d}{dx^\nu} \delta \Phi - \frac{d\Phi}{dx^\varrho} \frac{d\,\delta x^\varrho}{dx^\nu}. \tag{6}$$

Eine Größe $\mathfrak{R}$ heißt eine skalare Dichte (in bezug auf die Gruppe), wenn sie folgende Transformationseigenschaft hat:

$$\delta^* \mathfrak{R} + \frac{d}{dx^\nu}(\mathfrak{R}\,\delta x^\nu) = 0, \tag{7}$$

oder auch nach (4)

$$\delta \mathfrak{R} + \mathfrak{R} \frac{d\,\delta x^\nu}{dx^\nu} = 0. \tag{8}$$

Größen hängen im allgemeinen von zweierlei Indizes ab: erstens von Indizes $\alpha, \beta, \gamma, \ldots$, deren Wertbereich derjenige vom Index α in Q_α ist, zweitens vom Indizes $\mu, \nu, \ldots$, die, wie der Index von x^ν, von 1 bis 4 laufen. Insbesondere vertritt der Index r von ξ^r ein oder mehrere Systeme von Indizes $(\alpha, \beta, \ldots; \mu, \nu, \ldots)$, die in einer beliebigen eindimensionalen Folge numeriert sind. Die Indizes von der Art $\alpha, \beta, \ldots$ können auch ihrerseits mehrfach sein und insbesondere Systeme von Indizes $\mu, \nu, \ldots$ enthalten.

Ein kontravarianter Tensor $K^{\alpha\nu}$ wird definiert durch die Transformationseigenschaft

$$\delta K^{\alpha\nu} = K^{\alpha\mu} \frac{d\,\delta x^\nu}{dx^\mu} - \underline{K^{\beta\nu} \frac{\partial\,\delta Q_\beta}{\partial Q_\alpha}}; \tag{9}$$

im Falle der q-Zahlen enthält diese Erklärung wegen des unterstrichenen Gliedes eine Willkür, welche wir durch die Festsetzung

$$\underline{K^{\beta\nu} \frac{\partial\,\delta Q_\beta}{\partial Q_\alpha}} = \frac{1}{2}\left(K^{\beta\nu} \frac{\partial\,\delta Q_\beta}{\partial Q_\alpha} + \frac{\partial\,\delta Q_\beta}{\partial Q_\alpha} K^{\beta\nu\dagger}\right)$$

beseitigen; dabei bezeichnet $x^\dagger$ das hermitisch konjugierte (adjungierte) von x. [Im folgenden gebrauchen wir allgemein die Bezeichnung

$$\underline{x} = \tfrac{1}{2}(x + x^\dagger).]$$

Durch diese Festsetzung bleibt ein hermitischer Tensor nach einer beliebigen Transformation der Gruppe hermitisch.

Ein kovarianter Tensor $K_{\alpha\nu}$ hat die Transformationseigenschaft:

$$(10) \qquad \delta K_{\alpha\nu} = - K_{\alpha\mu} \frac{d\,\delta x^\mu}{dx^\nu} + K_{\beta\nu} \underline{\frac{\partial\,\delta Q_\alpha}{\partial Q_\beta}};$$

analog zu (9) und (10) wird die Variation des gemischten Tensors $K_{\alpha\beta\ldots}{}^{\gamma\delta\ldots}{}_{\mu\nu\ldots}{}^{\pi\varrho\ldots}$ gebildet.

Eine Tensordichte $\mathfrak{K}^{\alpha\nu}$ transformiert sich wie das Produkt eines Tensors $K^{\alpha\nu}$ mit einer skalaren Dichte $\mathfrak{K}$, also:

$$(11) \qquad \delta\,\mathfrak{K}^{\alpha\nu} = \mathfrak{K}^{\alpha\mu} \frac{d\,\delta x^\nu}{dx^\mu} - \mathfrak{K}^{\beta\nu} \underline{\frac{\partial\,\delta Q_\beta}{\partial Q_\alpha}} - \mathfrak{K}^{\alpha\nu} \frac{d\,\delta x^\mu}{dx^\mu}.$$

Jetzt sind wir imstande, die Invarianzforderung bezüglich der Lagrangefunktion $\mathfrak{L}$ zu formulieren. Damit nämlich das Integral $\int \mathfrak{L}\, dx^1\, dx^2\, dx^3\, dx^4$ invariant ist, soll nach bekannten Schlüssen [1]) $\mathfrak{L}$ bis auf eine Divergenz $\mathfrak{L}' \equiv \frac{d\,\mathfrak{K}^\nu}{dx^\nu}$ eine skalare Dichte sein. In Formeln:

$$(12) \qquad \delta(\mathfrak{L} + \mathfrak{L}') + (\mathfrak{L} + \mathfrak{L}') \frac{d\,\delta x^\nu}{dx^\nu} = 0.$$

Da es uns, wie gesagt, nicht auf die größte Allgemeinheit ankommt, wollen wir uns damit begnügen, der Reihe nach folgende chrakteristische Fälle zu behandeln:

1° $\mathfrak{L}' = 0$, d. h. $\mathfrak{L}$ ist selber eine skalare Dichte:

$$(13) \qquad \delta\,\mathfrak{L} + \mathfrak{L} \frac{d\,\delta x^\nu}{dx^\nu} = 0;$$

---

1) Vgl. etwa E. Noether, Gött. Nachr. 1918. S. 211. — Die Divergenz $\frac{d\,\mathfrak{K}^\nu}{d\,x^\nu}$ tritt dann auf, wenn das Integral $\int \mathfrak{L}\, dx^1 \ldots dx^4$ nicht bei *beliebigem* Integrationsgebiet invariant ist, sondern nur wenn die $\mathfrak{K}^\nu$ am Rande verschwinden.

2° $\mathfrak{L}'$ enthält die zweiten Ableitungen

$$Q_{\alpha,\nu\varrho} \equiv \frac{\partial^2 Q_\alpha}{\partial x^\nu \partial x^\varrho}$$

nur linear, d. h.

$$\mathfrak{L}' \equiv \frac{d}{dx^\nu}[\mathfrak{f}^{\nu,\alpha\varrho}(Q)\,Q_{\alpha,\varrho}] \tag{14}$$

und es ist $j = 0$ [vgl. Formel (3)].

In beiden Fällen zerfällt die Untersuchung in zwei Schritte: a) Durchführung des erweiterten Hamiltonschen Verfahrens; b) Beweis der Kovarianz desselben bezüglich der betrachteten Gruppe.

Wir beginnen mit dem ersten Falle.

§ 2. Die konjugierten Impulse und die Identitäten

Von jetzt an legen wir also die Forderung (13) zugrunde. Wir setzen zunächst

$$\mathfrak{P}^{\alpha\nu} = \frac{\partial \mathfrak{L}}{\partial Q_{\alpha,\nu}} \tag{15}$$

und nehmen als Impulse

$$\mathfrak{P}^\alpha \equiv \mathfrak{P}^{\alpha 4} = \frac{\partial \mathfrak{L}}{\partial \dot{Q}_\alpha}\,. \tag{16}$$

Beschränken wir uns zuerst auf die klassische (c-Zahl-) Theorie.

Ersetzen wir in (13) die δQ_α, $Q_{\alpha,\nu}$ und δx^ν durch ihre Werte (2), (6) als Funktionen der ξ^r und Ableitungen, so bekommen wir mehrere Identitäten, indem wir ausdrücken, daß die Koeffizienten der einzelnen Ableitungen von ξ^r identisch verschwinden sollen. Diese Identitäten enthalten aber im allgemeinen die $\dot{Q}_\alpha$ nicht nur durch die soeben eingeführten Funktionen $\mathfrak{P}^\alpha$, sondern auch in anderen Verbindungen (z. B. durch die anderen $\mathfrak{P}^{\alpha\nu}$, $\nu \neq 4$); für die Auflösung des Gleichungssystems (16) nach den $\dot{Q}_\alpha$ bieten sie also kein Interesse: sie stellen einfach Beziehungen dar, die jede Lösung $\dot{Q}_\alpha(Q, \mathfrak{P})$ dieses Systems von selbst erfülllt. Wesentlich anders liegen aber die Verhältnisse, wenn einige der betrachteten Identitäten nur die Q_α (nebst räumlichen Ableitungen) und die $\mathfrak{P}^\alpha$ enthalten: sie bedeuten dann, daß die Gleichungen (16) nicht alle voneinander unabhängig sind, so daß die allgemeine Lösung von gewissen willkürlichen Parametern (genauer: Raumzeitfunktionen) abhängt.

Nun tritt der letztere Fall bei der Gruppe (2) immer auf. Die höchsten in (13) vorkommenden Ableitungen von ξ^r sind die

$$\frac{\partial^{j+1}\,\xi^r}{\partial x^\sigma \cdots \partial x^\tau\, dx^\nu};$$

nach der Voraussetzung (3) lauten die entsprechenden Identitäten

$$\text{(17c)[1]} \qquad \sum \mathfrak{P}^{a\nu}\, c_{ar}^{\sigma\ldots\tau} \equiv 0,$$

wobei die Summe sich über alle Permutationen der Zahlen $\nu, \sigma, \ldots, \tau$ erstreckt. Für $\nu = \sigma = \ldots \tau = 4$ hat man insbesondere

$$\text{(18c)} \qquad \mathfrak{P}^a\, c_{ar}^{44\ldots 4} \equiv 0:$$

da nun die c nur die Q_a enthalten, haben wir in (18c) r_0 Identitäten von der zuletzt besprochenen Form vor uns, welche wir „eigentliche" Identitäten nennen wollen. Es ist ferner leicht einzusehen, daß im allgemeinen (d. h. falls die Lagrangefunktion keine spezielleren Eigenschaften besitzt) keine weiteren eigentlichen Identitäten vorkommen. Die allgemeinste Lösung $\dot{Q}_a$ $(Q, \mathfrak{P}, \lambda)$ von (16) hängt also von r_0 willkürlichen Parametern λ ab.

In den bisherigen, in der Einleitung erwähnten Methoden half man sich entweder durch Zerstören der Invarianzeigenschaft der Lagrangefunktion (1. und 3. Methode) oder durch Auszeichnung einer speziellen Lösung $\dot{Q}_a\,(Q, \mathfrak{P}, \lambda^0)$ (2. Methode). Im Gegensatz dazu ist der Grundgedanke der neuen Methode der, die Hamiltonsche Funktion in der üblichen Weise mittels der allgemeinen Lösung $\dot{Q}_a(Q, \mathfrak{P}, \lambda)$ mit unbestimmten λ^r zu konstruieren, ohne sich zunächst um die eigentlichen Identitäten zu kümmern: Feldgleichungen und Vertauschungsrelationen haben die kanonische Form, die ersteren enthalten die λ^r. Zu diesem kanonischen Schema kommen schließlich die eigentlichen Identitäten als Nebenbedingungen hinzu. Wir werden

1) Den Nummern der Formeln, die nur für c-Zahlen unbeschränkte Gültigkeit besitzen, wird der Buchstabe c angehängt.

sehen, daß diese Methode außer ihrer Einfachheit noch den großen Vorteil hat, daß der Kovarianzbeweis des Verfahrens ohne Schwierigkeit durchführbar ist.

§ 3. Übergang zu den q-Zahlen

Zuvor müssen wir untersuchen, wie sich beim Übergang zu den q-Zahlen die eben geschilderten Verhältnisse gestalten. Nach (1) lautet dann (15):

$$\mathfrak{P}^{\alpha\nu} = \tfrac{1}{2}(\mathfrak{p}^{\alpha\nu} + \mathfrak{p}^{\alpha\nu\dagger}) = \underline{\mathfrak{p}^{\alpha\nu}}, \tag{19}$$

mit

$$\mathfrak{p}^{\alpha\nu} = \mathfrak{A}^{\alpha\nu;\,\beta\mu} Q_{\beta,\mu} + \mathfrak{B}^{\alpha\nu}. \tag{20}$$

Durch einen Strich über einen Index von der Art $\mu : \bar{\mu}$ deuten wir an, daß er nur von 1 bis 3 läuft; für überstrichene Indizes soll die Regel vom Weglassen des Summenzeichens ebenfalls gelten. Mit dieser Bezeichnung schreiben wir nach (19) und (20)

$$\left\{\begin{aligned} \mathfrak{P}^{\alpha} &= \underline{\mathfrak{p}^{\alpha}}, \\ \mathfrak{p}^{\alpha} &= \mathfrak{A}^{\alpha\beta}\dot{Q}_{\beta} + \mathfrak{D}^{\alpha}, \end{aligned}\right. \tag{21}$$

wobei

$$\left\{\begin{aligned} \mathfrak{A}^{\alpha\beta} &\equiv \mathfrak{A}^{\alpha 4;\,\beta 4} \\ \mathfrak{D}^{\alpha} &\equiv \mathfrak{A}^{\alpha 4;\,\beta\bar{\mu}} Q_{\beta,\bar{\mu}} + \mathfrak{B}^{\alpha 4} \end{aligned}\right. \tag{22}$$

gesetzt ist. Es ist in diesen Formeln

$$\mathfrak{A}^{\alpha\nu;\,\beta\mu} = \mathfrak{A}^{\beta\mu;\,\alpha\nu},$$

insbesondere

$$\mathfrak{A}^{\alpha\beta} = \mathfrak{A}^{\beta\alpha},$$

angenommen worden, was natürlich keine Einschränkung bedeutet.

Die Überlegung des vorigen Paragraphen liefert jetzt statt (17c) und (18c)

$$\underline{\sum c^{\sigma\dots\tau}_{\alpha r}\,\mathfrak{p}^{\alpha\nu}} = 0 \tag{23}$$

und

$$\underline{c^{4\dots 4}_{\alpha r}\;\mathfrak{p}^{\alpha}} = 0. \tag{24}$$

Da insbesondere (24) in den $\dot{Q}_{\alpha}$ identisch gilt, so ist nach (21)

$$c^{4\dots 4}_{\alpha r}\;\mathfrak{A}^{\alpha\beta} = 0, \tag{25}$$

$$c^{4\dots 4}_{\alpha r}\;\mathfrak{D}^{\alpha} = 0: \tag{26}$$

die Koeffizienten der Lagrangefunktion müssen u. a. diese Beziehungen erfüllen, damit $\mathfrak{L}$ die verlangte Dichteeigenschaft haben kann. Die Relationen (25) und (26) heben wir für späteren Gebrauch hervor.

Nun aber können wir nicht weiterkommen, ohne etwas über die Vertauschungsrelationen $[Q_\alpha, \dot{Q}_\beta]$ zu wissen. Wenn wir die $\dot{Q}_\beta$ als Funktionen der Q_α und $\mathfrak{P}^\alpha$ kennen würden, so könnten wir den Wert von $[Q_\alpha, \dot{Q}_\beta]$ aus den kanonischen Vertauschungsrelationen, die wir, wie gesagt, beizubehalten wünschen, ableiten. Es ist indessen nicht einmal von vornherein sicher, ob wir aus (21) die $\dot{Q}_\beta$ als Funktionen der *Matrizen* $\mathfrak{P}^\alpha$ ableiten können, oder nur als Funktionen der *Matrixelemente* von $\mathfrak{P}^\alpha$. Der einzige Ausweg ist der, daß wir versuchsweise eine *Annahme* über $[Q_\alpha, \dot{Q}_\beta]$ machen, auf Grund deren die Lösung von (21) die Gestalt $\dot{Q}_\alpha(Q, \mathfrak{P}, \lambda)$ annimmt und nachher prüfen, ob die gemachte Annahme mit den kanonischen Vertauschungsrelationen verträglich ist.

Eine naheliegende Annahme ist folgende: die $[Q_\alpha, \dot{Q}_\beta]$ sollen *schiefe* Funktionen[1]) von den Q_α und $Q_{\alpha,\bar{\nu}}$, nicht aber von den $\dot{Q}_\alpha$ (bzw. den $\mathfrak{P}^\alpha$) sein. (Ob dabei, wenn Q_α und $\dot{Q}_\beta$ im selben Punkt genommen sind, unbestimmte Faktoren, wie $\delta(0)$, vorkommen, ist gleichgültig). Führen wir einige unmittelbare Folgerungen dieser Annahme an:

1. Nach (20) sind ebenfalls die $[Q_\alpha, \mathfrak{p}^{\beta\nu}]$ und $[Q_\alpha, \mathfrak{p}^{\beta\nu\dagger}]$ schiefe Funktionen der Q_α und $Q_{\alpha,\bar{\nu}}$ allein.

2. Die $[Q_\alpha, \dot{Q}_\beta]$, $[Q_\alpha, \mathfrak{p}^{\beta\nu}]$ und $[Q_\alpha, \mathfrak{p}^{\beta\nu\dagger}]$ sind mit jeder Funktion der Q_α und $Q_{\alpha,\bar{\nu}}$ vertauschbar.

3. Es ist

$$[Q_\alpha, \mathfrak{p}^{\beta\nu}] = [Q_\alpha, \mathfrak{p}^{\beta\nu\dagger}]. \tag{27}$$

Infolgedessen kann man statt (23) und (24)

$$\sum \underline{c_{\alpha r}^{\sigma \ldots \tau}\, \mathfrak{P}^{\alpha\nu}} = 0, \tag{28}$$

$$\mathfrak{F}_r \equiv \underline{c_{\alpha r}^{4 \ldots 4}\, \mathfrak{P}^\alpha} = 0 \tag{29}$$

schreiben.

Aus (25) folgt nun, daß die N linearen Gleichungen

$$\mathfrak{A}^{\alpha\beta}\dot{Q}_\beta + \dot{Q}_\beta\mathfrak{A}^{\beta\alpha} = 2(\mathfrak{P}^\alpha - \mathfrak{D}^\alpha) \tag{21}$$

1) Eine q-Zahl x heißt *schief*, wenn $x^\dagger = -x$.

nicht alle unabhängig sind, sondern daß ihre Determinante $|\mathfrak{A}^{\alpha\beta}|$ den Rang $N-r_0$ hat. Da sie symmetrisch ist, gibt es einen von Null verschiedenen Hauptminor vom Grade $N-r_0$; die dazu bezüglichen Indizes wollen wir mit einem Strich versehen:

$$|\mathfrak{A}^{\alpha'\beta'}| \neq 0,$$

während die übrigen doppelt gestrichen seien: α'', β'', $\cdots$. Die Determinante $|\mathfrak{A}^{\alpha'\beta'}|$, sowie ihre reziproke $|\mathfrak{A}_{\alpha'\beta'}|$, sind symmetrisch und es gilt:

$$\mathfrak{A}^{\alpha'\beta'}\mathfrak{A}_{\beta'\gamma'} = \delta^{\alpha'}_{\gamma'}, \tag{30}$$

wobei δ^{α}_{β} wie üblich gleich 0 oder 1 ist, je nachdem $\alpha \neq \beta$ oder $\alpha = \beta$.

Gelingt es also, eine *spezielle* Lösung $\dot{Q}_\beta{}^0(Q, \mathfrak{P})$ von (21) zu finden, so hat die allgemeinste Lösung die Form:

$$\dot{Q}_\beta = \dot{Q}_\beta{}^0 + \lambda^r x_{\beta r},$$

wo die λ^r r_0 willkürliche Parameter und $x_{\beta r}$ r_0 unabhängige Lösungen der homogenen Gleichungen

$$\mathfrak{A}^{\alpha\beta} x_{\beta r} + x_{\beta r}\mathfrak{A}^{\beta\alpha} = 0$$

darstellen. Nach (25) können wir nun

$$x_{\beta r} = c^{4\cdots 4}_{\beta r}$$

wählen und schreiben:

$$\dot{Q}_\beta = \dot{Q}_\beta{}^0 + \lambda^r\, c^{4\cdots 4}_{\beta r}. \tag{31}$$

Ferner behaupte ich, daß

$$\begin{cases} \dot{Q}^0_{\beta'} = \frac{1}{2}\{\mathfrak{A}_{\beta'\gamma'}(\mathfrak{P}^{\gamma'} - \mathfrak{D}^{\gamma'}) + (\mathfrak{P}^{\gamma'} - \mathfrak{D}^{\gamma'})\mathfrak{A}_{\gamma'\beta'}\} \\ \dot{Q}^0_{\beta''} = 0 \end{cases} \tag{32}$$

eine spezielle Lösung von (21) ist; ist dies nachgewiesen, so ist uns die Auflösung von (21) nach den $\dot{Q}_\beta$ wirklich gelungen: denn die Lösung (31) hat offenbar die verlangte Eigenschaft, daß vermöge der kanonischen Vertauschungsrelationen $[Q_\alpha, \dot{Q}_\beta]$ eine schiefe Funktion der Q_α und $Q_{\alpha,\bar{\nu}}$ wird.

Durch Einsetzen von (32) in die linke Seite von (21), die wir für einen Augenblick $\mathfrak{T}_\alpha$ nennen wollen, bekommt man

$$
\begin{aligned}
\mathfrak{T}^{\alpha} &= \tfrac{1}{2}\,\mathfrak{A}^{\alpha\beta'}\{\mathfrak{A}_{\beta'\gamma'}(\mathfrak{P}^{\gamma'}-\mathfrak{D}^{\gamma'})+(\mathfrak{P}^{\gamma'}-\mathfrak{D}^{\gamma'})\,\mathfrak{A}_{\gamma'\beta'}\}\\
&\quad+\tfrac{1}{2}\{\mathfrak{A}_{\beta'\gamma'}(\mathfrak{P}^{\gamma'}-\mathfrak{D}^{\gamma'})+(\mathfrak{P}^{\gamma'}-\mathfrak{D}^{\gamma'})\,\mathfrak{A}_{\gamma'\beta'}\}\,\mathfrak{A}^{\beta'\alpha}\\
&= \mathfrak{A}^{\alpha\beta'}\mathfrak{A}_{\beta'\gamma'}(\mathfrak{P}^{\gamma'}-\mathfrak{D}^{\gamma'})+(\mathfrak{P}^{\gamma'}-\mathfrak{D}^{\gamma'})\,\mathfrak{A}_{\gamma'\beta'}\mathfrak{A}^{\beta'\alpha}\\
&\quad+\tfrac{1}{2}\,\mathfrak{A}^{\alpha\beta'}[\mathfrak{P}^{\gamma'}-\mathfrak{D}^{\gamma'},\ \mathfrak{A}_{\gamma'\beta'}]+\tfrac{1}{2}[\mathfrak{A}_{\beta'\gamma'},\ \mathfrak{P}^{\gamma'}-\mathfrak{D}^{\gamma'}]\,\mathfrak{A}^{\beta'\alpha}\\
&= \mathfrak{A}^{\alpha\beta'}\mathfrak{A}_{\beta'\gamma'}(\mathfrak{P}^{\gamma'}-\mathfrak{D}^{\gamma'})+(\mathfrak{P}^{\gamma'}-\mathfrak{D}^{\gamma'})\,\mathfrak{A}_{\gamma'\beta'}\mathfrak{A}^{\beta'\alpha},
\end{aligned}
$$

wegen der Folgerung 2 aus unserer Annahme. Für $\alpha=\alpha'$ ist bereits nach (30)

$$\mathfrak{T}^{\alpha'} = 2\,(\mathfrak{P}^{\alpha'}-\mathfrak{D}^{\alpha'}).$$

Nun sind nach der Theorie der linearen Gleichungen und unter Benutzung unserer Annahme über $[Q_\alpha, \dot{Q}_\beta]$ die Identitäten (29) äquivalent mit

$$\mathfrak{P}^{\alpha''} = \underline{\mathfrak{A}^{\alpha''\beta'}\mathfrak{A}_{\beta'\gamma'}\mathfrak{P}^{\gamma'}}$$

und ebenso (26) äquivalent mit

$$\mathfrak{D}^{\alpha''} = \mathfrak{A}^{\alpha''\beta'}\mathfrak{A}_{\beta'\gamma'}\mathfrak{D}^{\gamma'};$$

folglich ist auch

$$\mathfrak{T}^{\alpha''} = 2\,(\mathfrak{P}^{\alpha''}-\mathfrak{D}^{\alpha''}),$$

womit der Nachweis, daß (31), (32) die allgemeinste Lösung von (21) im Einklang mit den kanonischen Vertauschungsrelationen darstellt, vollständig erbracht ist.

§ 4. Aufstellung der Hamiltonfunktion

Klassisch lautet die Hamiltonfunktion

$$\mathfrak{H} = \mathfrak{P}^{\alpha}\dot{Q}_{\alpha} - \mathfrak{L};$$

von jedem quantenmechanischen Ansatz müssen wir nun verlangen, daß

$$\frac{\partial\,\mathfrak{H}}{\partial\,Q_{\alpha,\bar{\nu}}} = -\frac{\partial\,\mathfrak{L}}{\partial\,Q_{\alpha,\bar{\nu}}}, \tag{33}$$

eine Eigenschaft, die sich für die Durchführung der Theorie als unentbehrlich erweisen wird.

Nun ist

$$\left(\frac{\partial\,\mathfrak{L}}{\partial\,Q_{\alpha,\bar{\nu}}}\right)_{\mathfrak{P}^{\alpha}} = \left(\frac{\partial\,\mathfrak{L}}{\partial\,Q_{\alpha,\bar{\nu}}}\right)_{\dot{Q}_{\alpha}} + \underline{\left(\frac{\partial\,\dot{Q}_{\beta}}{\partial\,Q_{\alpha,\bar{\nu}}}\right)_{\mathfrak{P}^{\alpha}}\cdot\mathfrak{p}^{\beta}}$$

und nach (31), (32) enthält $\left(\frac{\partial \dot{Q}_\beta}{\partial Q_{\alpha,\bar{\nu}}}\right)_{\mathfrak{P}^\alpha}$ nicht mehr die $\mathfrak{P}^\alpha$ folglich dürfen wir schreiben:

$$\left(\frac{\partial \mathfrak{L}}{\partial Q_{\alpha,\bar{\nu}}}\right)_{\mathfrak{P}^\alpha} = \left(\frac{\partial \mathfrak{L}}{\partial Q_{\alpha,\bar{\nu}}}\right)_{\dot{Q}_\alpha} + \underline{\left(\frac{\partial \dot{Q}_\beta}{\partial Q_{\alpha,\bar{\nu}}}\right)_{\mathfrak{P}^\alpha} \cdot \mathfrak{P}^\beta}\,.$$

Die gewünschte Eigenschaft (33) hat also der Ansatz

$$\mathfrak{H} = \underline{\dot{Q}_\alpha \mathfrak{P}^\alpha} - \mathfrak{L}\,. \tag{34}$$

Da nach (25) und (26)

$$\mathfrak{L}[Q;\, \dot{Q}(Q, \mathfrak{P}, \lambda)] = \mathfrak{L}(Q, \dot{Q}^0)$$

ist, können wir schreiben gemäß der Bezeichnung (29)

$$\mathfrak{H} = \mathfrak{H}_0 + \lambda^r \mathfrak{F}_r\,, \tag{35}$$

mit

$$\mathfrak{H}_0 = \underline{\dot{Q}_\alpha{}^0 \mathfrak{P}^\alpha} - \mathfrak{L}\,[Q, \dot{Q}^0(Q, \mathfrak{P})]\,. \tag{36}$$

Nun setzen wir die kanonischen Vertauschungsrelationen an

$$\begin{cases} [Q_\alpha(\mathfrak{r}), Q_\beta(\mathfrak{r}')] = [\mathfrak{P}^\alpha(\mathfrak{r}), \mathfrak{P}^\beta(\mathfrak{r}')] = 0\,, \\ [\mathfrak{P}^\alpha(\mathfrak{r}), Q_\beta(\mathfrak{r}')] = \omega\, \delta_\beta{}^\alpha\, \delta(\mathfrak{r} - \mathfrak{r}'),\ \omega = \frac{h c}{2 \pi i}\,, \end{cases} \tag{37}$$

sowie die Feldgleichungen

$$\begin{cases} [\overline{\mathfrak{H}}, Q_\alpha] = \omega\, \dot{Q}_\alpha \\ [\overline{\mathfrak{H}}, \mathfrak{P}^\alpha] = \omega\, \dot{\mathfrak{P}}^\alpha\,, \end{cases} \tag{38}$$

wobei die Bezeichnung

$$\overline{\mathfrak{A}} \equiv \int \mathfrak{A}\, d x^1 d x^2 d x^3 \tag{39}$$

gebraucht ist; das Integrationsgebiet muß so gewählt werden, daß die Feldgrößen am Rande *konstante* Werte annehmen und zwar solche, daß $\mathfrak{L}$ dort verschwindet.

Zu (37) und (38) kommen noch als Nebenbedingungen die eigentlichen Identitäten (29) $\mathfrak{F}_r = 0$ hinzu. Aber es muß bewiesen werden, daß es erlaubt ist, die q-Zahlen $\mathfrak{F}_r$ alle gleichzeitig gleich Null zu setzen; mit anderen Worten, daß die $\mathfrak{F}_r$ untereinander vertauschbar sind, wenigstens auf Grund der Nebenbedingungen $\mathfrak{F}_r = 0$ selber.

Die jetzt folgenden Betrachtungen dienen nicht nur zu diesem Zwecke, sondern sind auch für den später zu erbringenden Kovarianzbeweis grundlegend.

Wir definieren zunächst den Impuls-Energie-Pseudotensor[1])

(40) $$\mathfrak{G}_\mu^\nu = \underline{\mathfrak{P}^{\alpha\nu} Q_{\alpha,\mu}} - \delta_\mu^\nu \mathfrak{L},$$

sodann die Impuls-Energie-Pseudodichte

(41) $$\mathfrak{G}_\mu \equiv \mathfrak{G}_\mu^4 = \underline{\mathfrak{P}^\alpha Q_{\alpha,\mu}} - \delta_\mu^4 \mathfrak{L},$$

deren vierte Pseudokomponente die Hamiltonfunktion (34) ist:

$$\mathfrak{H} \equiv \mathfrak{G}_4 \equiv \mathfrak{G}_4^4.$$

Die Komponenten des Gesamtimpulses sind dann $\overline{\mathfrak{G}_{\bar\nu}}$, die Gesamtenergie $\overline{\mathfrak{H}}$.

Die V.-R. (Vertauschungsrelationen) von $\overline{\mathfrak{H}}$ mit den Q_α, $\mathfrak{P}^\alpha$ sind durch (38) gegeben. Was die $\overline{\mathfrak{G}_{\bar\nu}}$ betrifft, so finden wir zunächst auf Grund von (37)

(42) $$\begin{cases} [\mathfrak{G}_{\bar\nu}(\mathfrak{r}), Q_\alpha(\mathfrak{r}')] = \omega \dfrac{\partial Q_\alpha}{\partial x^{\bar\nu}} \delta(\mathfrak{r}-\mathfrak{r}'), \\ [\mathfrak{G}_{\bar\nu}(\mathfrak{r}), \mathfrak{P}^\alpha(\mathfrak{r}')] = -\omega \mathfrak{P}^\alpha \dfrac{\partial \delta(\mathfrak{r}-\mathfrak{r}')}{\partial x^{\bar\nu}}, \end{cases}$$

sodann

$$[\overline{\mathfrak{G}}_{\bar\nu}, \Phi(Q, \mathfrak{P})] = \omega \frac{d\Phi}{dx^{\bar\nu}},$$

allgemeiner also

(43) $$\omega \frac{d\Phi}{dx^\nu} = [\overline{\mathfrak{G}}_\nu, \Phi(Q, \mathfrak{P}, x)] + \omega \frac{\partial \Phi}{\partial x^\nu}.$$

Daraus folgt unmittelbar

(44) $$[\overline{\mathfrak{G}}_\nu, \overline{\mathfrak{G}}_\mu] = 0:$$

ein Ausdruck für die Vertauschbarkeit der Differentiationen $\dfrac{d}{dx^\nu}$, dessen physikalischer Inhalt in der zeitlichen Konstanz der $\overline{\mathfrak{G}}_\nu$ als Folge der Gleichungen (38), (37) besteht.[2])

1) Der Vorsatz „Pseudo" deutet an, daß die betreffenden Größen keine Tensoren sind.

2) Falls die $\lambda^r x^4$ explizite enthalten, gilt (44) erst auf Grund der Nebenbedingungen (29).

§ 5. Quantenmechanischer Ausdruck der infinitesimalen Transformation der Gruppe

In diesem Paragraphen beweisen wir den Satz:

$$\omega\,\delta^* \Phi(Q, \mathfrak{P}) = [\overline{\mathfrak{M}}, \Phi], \tag{45}$$

wobei

$$\mathfrak{M} = \underline{\mathfrak{P}^\alpha \,\delta Q_\alpha} - \mathfrak{G}_\mu \,\delta x^\mu . \tag{46}$$

Dies soll auf Grund der Feldgleichungen (38) und der V.-R. (37) gelten, unter der Voraussetzung (13), daß $\mathfrak{L}$ eine skalare Dichte ist.

Um diesen Satz zu beweisen, genügt es zu zeigen, daß

$$\left\{\begin{aligned} \omega\,\delta^* Q_\alpha &= [\overline{\mathfrak{M}}, Q_\alpha], \\ \omega\,\delta^* \mathfrak{P}^\alpha &= [\overline{\mathfrak{M}}, \mathfrak{P}^\alpha]. \end{aligned}\right. \tag{47}$$

Nach (37) und (42) ist, wenn man bedenkt, daß δQ_α nach (2) nur die Q_α (nicht die $\mathfrak{P}^\alpha$) enthält,

$$[\overline{\mathfrak{M}}, Q_\alpha] = \omega\,\delta Q_\alpha - \frac{d\,Q_\alpha}{d\,x^{\bar\mu}}\,\delta x^{\bar\mu} - [\overline{\delta x^4 \cdot \mathfrak{H}}, Q_\alpha].$$

Nun ist, nach H. P. I, Formel (20),

$$\left\{\begin{aligned} [\overline{\delta x^4 \mathfrak{H}}, Q_\alpha] &= \omega \frac{\partial(\delta x^4 \mathfrak{H})}{\partial \mathfrak{P}^\alpha} = \delta x^4 \cdot [\overline{\mathfrak{H}}, Q_\alpha], \\ [\overline{\delta x^4 \mathfrak{H}}, \mathfrak{P}^\alpha] &= -\,\omega \left\{ \frac{\partial(\delta x^4 \mathfrak{H})}{\partial Q_\alpha} - \frac{d}{d\,x^{\bar\nu}}\,\frac{\partial(\delta x^4 \mathfrak{H})}{\partial Q_{\alpha,\bar\nu}} \right\} \\ &= \delta x^4 [\overline{\mathfrak{H}}, \mathfrak{P}^\alpha] + \omega \frac{d\,\delta x^4}{d\,x^{\bar\nu}}\,\frac{\partial \mathfrak{H}}{\partial Q_{\alpha,\bar\nu}}. \end{aligned}\right. \tag{48}$$

Folglich gilt tatsächlich, mit Rücksicht auf (4), die erste Formel (47), wenn man noch die erste Feldgleichung (38) benutzt.

Analog findet man unter Berücksichtigung der zweiten Formel (48), der zweiten Feldgleichung (38) und der Formel (33)

$$\left\{\begin{aligned} \frac{1}{\omega}[\overline{\mathfrak{M}}, \mathfrak{P}^\alpha] &= -\,\underline{\mathfrak{P}^\beta \frac{\partial\,\delta Q_\beta}{\partial Q_\alpha}} - \frac{d}{dx^{\bar\nu}}(\mathfrak{P}^\alpha \,\delta x^{\bar\nu}) \\ &\qquad - \frac{d\,\mathfrak{P}^\alpha}{dx^4}\,\delta x^4 + \mathfrak{P}^{\alpha\bar\nu}\,\frac{d\,\delta x^4}{dx^{\bar\nu}} \\ &= -\,\underline{\mathfrak{P}^\beta \frac{\partial\,\delta Q_\beta}{\partial Q_\alpha}} + \mathfrak{P}^{\alpha\nu}\,\frac{d\,\delta x^4}{dx^\nu} - \frac{d}{dx^\nu}(\mathfrak{P}^\alpha \,\delta x^\nu). \end{aligned}\right. \tag{49}$$

Es bleibt nur noch übrig zu zeigen, daß die rechte Seite von (49) gleich $\delta^* \mathfrak{P}^\alpha$ ist. Berechnen wir also direkt $\delta \mathfrak{P}^\alpha$, oder vielmehr allgemeiner $\delta \mathfrak{P}^{\alpha\nu}$. Zunächst gilt:

$$\delta \mathfrak{P}^{\alpha\nu} = \delta \left(\frac{\partial \mathfrak{L}}{\partial Q_{\alpha,\nu}} \right) = \frac{\partial (\delta \mathfrak{L})}{\partial Q_{\alpha,\nu}} - \underline{\frac{\partial \mathfrak{L}}{\partial Q_{\beta,\mu}} \frac{\partial \delta Q_{\beta,\mu}}{\partial Q_{\alpha,\nu}}}, \tag{50}$$

und zwar bei c-Zahlen allgemein, bei q-Zahlen jedenfalls, wenn $\mathfrak{L}$ die Form (1) hat und $\dfrac{\partial \delta Q_{\beta,\mu}}{\partial Q_{\alpha,\nu}}$ die $\dot{Q}_\alpha$ bzw. $\mathfrak{P}^\alpha$ nicht enthält. Daß letzteres in unserem Falle zutrifft, zeigt die Formel (6), welche ergibt:

$$\begin{aligned} \frac{\partial}{\partial Q_{\alpha,\nu}} \delta Q_{\beta,\mu} &= \frac{\partial}{\partial Q_{\alpha,\nu}} \left\{ \frac{d}{dx^\mu} \delta Q_\beta - Q_{\beta,\varrho} \frac{d\,\delta x^\varrho}{dx^\mu} \right\} \\ &= \frac{\partial \delta Q_\beta}{\partial Q_\alpha} \delta_\mu^\nu - \frac{d\,\delta x^\nu}{dx^\mu} \delta_\beta^\alpha . \end{aligned}$$

Dies, in (50) eingesetzt, liefert

$$\delta \mathfrak{P}^{\alpha\nu} = \underline{- \mathfrak{P}^{\beta\nu} \frac{\partial \delta Q_\beta}{\partial Q_\alpha}} + \mathfrak{P}^{\alpha\mu} \frac{d\,\delta x^\nu}{dx^\mu} + \frac{\partial (\delta \mathfrak{L})}{\partial Q_{\alpha,\nu}};$$

benutzen wir jetzt (13), so kommt

$$\delta \mathfrak{P}^{\alpha\nu} = \underline{- \mathfrak{P}^{\beta\nu} \frac{\partial \delta Q_\beta}{\partial Q_\alpha}} + \mathfrak{P}^{\alpha\mu} \frac{d\,\delta x^\nu}{dx^\mu} - \mathfrak{P}^{\alpha\nu} \frac{d\,\delta x^\mu}{dx^\mu}; \tag{51}$$

d. h. wie der Vergleich mit (11) lehrt: $\mathfrak{P}^{\alpha\nu}$ ist eine Tensordichte. Aus (51) folgt nun sofort mit Rücksicht auf (4) für $\delta^* \mathfrak{P}^\alpha \equiv \delta^* \mathfrak{P}^{\alpha 4}$ der Ausdruck (49).

Somit ist die Formel (45) bewiesen.

§ 6. Die $\overline{\mathfrak{F}_r}$ als spezielle infinitesimale Transformationen

Betrachten wir einen bestimmten, aber beliebigen Schnitt $x^4 = x_0^4$. Auf diesem Schnitt betrachten wir die Transformationen unserer Gruppe (2), welche durch die Forderungen

$$\left\{ \begin{aligned} & (\xi^r)_{x^4 = x_0^4} = \left(\frac{\partial \xi^r}{\partial x^\sigma} \right)_{x^4 = x_0^4} = \cdots = \left(\frac{\partial^{j-1} \xi^r}{\partial x^\sigma \cdots \partial x^\tau} \right)_{x^4 = x_0^4} = 0, \\ & \left(\frac{\partial^j \xi^r}{\partial x^\sigma \cdots \partial x^\tau} \right)_{x^4 = x_0^4} = 0, \text{ wenn nicht alle } \sigma, \ldots, \tau \text{ gleich 4 sind,} \\ & \left[\frac{\partial^j \xi^r}{(\partial x^4)^j} \right]_{x^4 = x_0^4} = \varepsilon^r \end{aligned} \right. \tag{52}$$

definiert sind, wobei die ε^r beliebige Raumfunktionen sind.

Wegen der Voraussetzung (3) führen diese Transformationen nicht aus dem Schnitt $x^4 = x_0^{\ 4}$ heraus. Sie bilden in jedem Punkte dieses Schnittes eine endliche kontinuierliche Untergruppe der Gruppe (2), deren infinitesimale Transformation nach (45) und (46) gegeben ist durch

$$\omega\, \delta' \Phi(Q, \mathfrak{P}) = [\overline{\varepsilon^r \mathfrak{F}_r}, \Phi].$$

(Hierin sind Q, $\mathfrak{P}$, $\mathfrak{F}_r$ für $x^4 = x_0^{\ 4}$ zu nehmen.)

Der zweite Fundamentalsatz von Lie über endliche Transformationsgruppen besagt, angewandt auf diese Untergruppe, daß in jedem Punkte des Schnittes

$$[\mathfrak{F}_r, [\mathfrak{F}_s, \Phi]] - [\mathfrak{F}_s, [\mathfrak{F}_r, \Phi]] = c^t_{rs} [\mathfrak{F}_t, \Phi]$$

gilt, wo die c^t_{rs} die vom Punkte $(x^1, x^2, x^3, x_0^{\ 4})$ abhängigen „Strukturkonstanten" der Gruppe sind. Nach der Jacobischen Identität über Klammersymbole wird die linke Seite einfach gleich

$$[[\mathfrak{F}_s, \mathfrak{F}_r], \Phi];$$

also bekommen wir

$$[\mathfrak{F}_s, \mathfrak{F}_r] = c^t_{rs}\, \mathfrak{F}_t. \tag{53}$$

Daraus folgt die zur Begründung des in § 4 dargelegten Verfahrens noch nötige Tatsache, daß *auf Grund von* $\mathfrak{F}_r = 0$ die $\mathfrak{F}_r$ untereinander vertauschbar sind.

§ 7. Die infinitesimale Transformation $\overline{\mathfrak{M}}$ als Integral der Bewegung

Kehren wir einen Augenblick zur reinen c-Zahltheorie zurück. Setzen wir

$$\mathfrak{M}^\nu = \underline{\mathfrak{P}^{\alpha\nu}\, \delta Q_\alpha} - \mathfrak{G}^\nu_\mu\, \delta x^\mu \tag{54}$$

und

$$\mathfrak{L}^\alpha = \frac{\partial \mathfrak{L}}{\partial Q_\alpha} - \frac{d}{dx^\nu}\, \frac{\partial \mathfrak{L}}{\partial Q_{\alpha,\nu}}, \tag{55}$$

so ist, wie leicht zu sehen, die Voraussetzung (13) gleichbedeutend mit

$$\frac{d\,\mathfrak{M}^\nu}{dx^\nu} + \mathfrak{L}^\alpha\, \delta^* Q_\alpha = 0; \tag{56c}$$

berücksichtigt man nun, daß nach (46) und (54)

$$\mathfrak{M} \equiv \mathfrak{M}^4$$

und entsprechend der Bedeutung der Bezeichnung (39)

$$\overline{\frac{d\,\mathfrak{M}^{\bar{\nu}}}{dx^{\bar{\nu}}}} = 0$$

ist, so folgt aus (56c)

$$\frac{d\,\overline{\mathfrak{M}}}{dx^4} = -\overline{\mathfrak{L}^\alpha\,\partial^* Q_\alpha}\,. \tag{57c}$$

Nun sind bekanntlich die Hamiltonschen Gleichungen (38) [vermöge der eigentlichen Identitäten (29)] äquivalent mit den Lagrangegleichungen

$$\mathfrak{L}^\alpha = 0\,.$$

Nach (57c) gilt somit, auf Grund von (13) und (38)

$$\frac{d\,\overline{\mathfrak{M}}}{dx^4} = 0\,. \tag{58}$$

Die Gleichung (56c) läßt sich nicht auf q-Zahlen übertragen. Die Ableitung von (58) gelingt jedoch unter Benutzung der nämlichen Voraussetzungen (13) und (38), bloß in etwas anderem Zusammenhang. Die beiden Relationen (13) und (38) wurden zur Ableitung der Formeln (43) und (45) wesentlich gebraucht. Wenden wir diese letzteren auf die Identität (5) an, wobei Φ nur von Q und $\mathfrak{P}$ abhängen möge:

$$[\overline{\mathfrak{M}}, [\overline{\mathfrak{G}}_\nu, \Phi]] = [\overline{\mathfrak{G}}_\nu, [\overline{\mathfrak{M}}, \Phi]] + \left[\omega\,\frac{\partial\,\overline{\mathfrak{M}}}{\partial x^\nu}, \Phi\right],$$

$$\left[[\overline{\mathfrak{G}}_\nu, \overline{\mathfrak{M}}] + \omega\,\frac{\partial\,\overline{\mathfrak{M}}}{\partial x^\nu}, \Phi\right] = 0$$

nach der Jacobischen Identität, oder schließlich

$$\left[\frac{d\,\overline{\mathfrak{M}}}{dx^\nu}, \Phi\right] = 0\,.$$

Insbesondere ist also $\frac{d\,\overline{\mathfrak{M}}}{dx^4}$ eine (eventuell von x^4 abhängige) c-Zahl. Nun kann diese c-Zahl, als Summe von lauter q-Zahlen, nichts anderes als Null sein. Diesen Schluß bestätigt übrigens die etwas mühsame Ausrechnung von $\frac{d\,\overline{\mathfrak{M}}}{dx^4}$.

Aus (58) lassen sich interessante Schlüsse ziehen über den Zusammenhang von $\overline{\mathfrak{M}}$ mit den Funktionen $\mathfrak{F}_r$. Durch partielle Integrationen wird $\overline{\mathfrak{M}}$ in die Form:

$$\overline{\mathfrak{M}} = \int d x^1 \, d x^2 \, d x^3 \sum_{i=0}^{i=j} \mathfrak{N}_r{}^i \, \frac{\partial^i \xi^r}{(\partial x^4)^i} \tag{59}$$

gebracht, wobei

$$\mathfrak{N}_r^j \equiv \mathfrak{F}_r \, . \tag{60}$$

(58) drückt sich dann folgendermaßen aus:

$$\int d x^1 \, d x^2 \, d x^3 \sum_{i=0}^{i=j} \mathfrak{N}_r{}^i \, \frac{\partial^{i+1} \xi^r}{(\partial x^4)^{i+1}}$$

$$= - \int d x^1 \, d x^2 \, d x^3 \sum_{i=0}^{i=j} \frac{d \, \mathfrak{N}_r{}^i}{d x^4} \, \frac{\partial^i \xi^r}{(\partial x^4)^i} \, .$$

Daraus folgert man durch Koeffizientenvergleich

$$\mathfrak{N}_r{}^i = - \frac{d \, \mathfrak{N}_r^{i+1}}{d x^4} \quad (i = 0, 1, \ldots, j-1) \tag{61}$$

und

$$\mathfrak{N}_r^j = 0 \, , \quad \frac{d \, \mathfrak{N}_r{}^0}{d x^4} = 0 \, . \tag{62}$$

Aus (60) und (61) ergibt sich

$$\mathfrak{N}_r{}^i = (-1)^{j-i} \, \frac{d^{j-i} \, \mathfrak{F}_r}{(d x^4)^{j-i}} \quad (i = 0, 1, \ldots, j), \tag{63}$$

also für $\overline{\mathfrak{M}}$ die merkwürdige Gestalt

$$\overline{\mathfrak{M}} = \int d x^1 \, d x^2 \, d x^3 \sum_{i=0}^{i=j} (-1)^{j-i} \, \frac{d^{j-i} \, \mathfrak{F}_r}{(d x^4)^{j-i}} \, \frac{\partial^i \xi^r}{(\partial x^4)^i} \, . \tag{63'}$$

Die erste Identität (62) ist nach (60) trivialerweise $\mathfrak{F}_r = 0$, die zweite aber besagt, daß auf Grund der Feldgleichungen und der Identitäten (29)

$$\frac{d^{j+1} \, \mathfrak{F}_r}{(d x^4)^{j+1}} = 0 \, . \tag{64}$$

Dies liefert die Antwort auf die Frage, inwieweit man durch abermalige Differentiation der Nebenbedingungen (29) neue Relationen bekommt.

Ist insbesondere $j = 1$, so sind die einzigen neuen Gleichungen $\frac{d\mathfrak{F}_r}{dx^4} = 0$, d. h.

$$[\overline{\mathfrak{H}}, \mathfrak{F}_r] + \omega \frac{\partial \mathfrak{F}_r}{\partial x^4} = 0 .$$

Ist nun die Lagrangefunktion von der Gestalt (1), d. h. gilt (35), so wird die letzte Gleichung mit Rücksicht auf (53)

$$[\overline{\mathfrak{H}}_0, \mathfrak{F}_r] + \omega \frac{\partial \mathfrak{F}_r}{\partial x^4} + c^t_{rs}\, \lambda^s \mathfrak{F}_t = 0$$

oder schließlich vermöge $\mathfrak{F}_r = 0$

$$[\overline{\mathfrak{H}}_0, \mathfrak{F}_r] + \omega \frac{\partial \mathfrak{F}_r}{\partial x^4} = 0 :$$

da weder die Nebenbedingungen noch diese neuen Gleichungen die λ^r enthalten, so bleiben dieselben wesentlich unbestimmt. (Anders aber, falls $j > 1$, denn bereits $\frac{d^2 \mathfrak{F}_r}{(dx^4)^2} = 0$ enthält die λ^r). Infolge der wesentlichen Unbestimmtheit der λ^r fehlen r_0 Feldgleichungen von der Form

$$\omega \dot{\mathfrak{P}}^a = [\overline{\mathfrak{H}}, \mathfrak{P}^a] ;$$

zum Ersatz reichen gerade die Gleichungen

$$\mathfrak{N}_r{}^0 \equiv \frac{d\mathfrak{F}_r}{dx^4} = 0 , \quad \text{d. h.} \quad [\overline{\mathfrak{H}}_0, \mathfrak{F}_r] + \omega \frac{\partial \mathfrak{F}_r}{\partial x^4} = 0$$

aus.

Im Falle $j = 0$ werden die fehlenden Feldgleichungen ersetzt durch die Identitäten $\mathfrak{F}_r = 0$ selber, die sich gemäß (64), d. h. $\frac{d\mathfrak{F}_r}{dx^4} = 0$ auf Grund der Feldgleichungen mit der Zeit fortpflanzen.

Eine letzte Bemerkung wollen wir noch an die Formel (63′) anknüpfen. Fragen wir nach der Untergruppe unserer Gruppe, die alle Punkte des Schnittes $x^4 = x_0{}^4$ invariant läßt: diese Untergruppe ist offenbar ein *Normalteiler*. Die Bedingungen

$$\delta x^\nu = 0 \quad \text{für} \quad x^4 = x_0{}^4$$

bedeuten

$$(\xi^r)_{x^4 = x_0^4} = \left(\frac{\partial \xi^r}{\partial x^\sigma}\right)_{x^4 = x_0^4} = \cdots = \left[\frac{\partial^k \xi^r}{\partial x^\sigma \cdots \partial x^\tau}\right]_{x^4 = x_0^4} = 0 ;$$

die Infinitesimaltransformation lautet demnach gemäß (63′)

$$\overline{\mathfrak{S}} = \int d\,x^1\,d\,x^2\,d\,x^3 \sum_{i=k+1}^{i=j} (-1)^{j-i} \frac{d^{j-i}\mathfrak{F}_r}{(d\,x^4)^{j-i}}\,\varepsilon_i^{\,r}, \tag{65}$$

wobei $\frac{d^{j-i}\mathfrak{F}_r}{(d\,x^4)^{j-i}}$ für $x^4 = x_0^{\,4}$ zu nehmen ist und die

$$\varepsilon_i^{\,r} \equiv \left[\frac{\partial^i\,\xi^r}{(\partial\,x^4)^i}\right]_{x^4 = x_0^4}$$

willkürliche Raumfunktionen sind. Die Gruppe $\overline{\mathfrak{S}}$ ist in jedem Punkte des Schnittes eine $r_0\,(j-k)$-parametrige invariante Untergruppe. Die im § 6 betrachtete Gruppe (52) ist eine Untergruppe davon.

§ 8. Kovarianz des Verfahrens gegenüber der Gruppe

Mittels der im Vorangehenden gewonnenen Ergebnisse sind wir nunmehr imstande, die Frage nach der Kovarianz des Verfahrens leicht zu erledigen.

Die Formel (45) besagt, daß bei einer beliebigen Transformation der Gruppe jedes Funktional $\Phi\,(Q, \mathfrak{P})$ einer unitären Ähnlichkeitstransformation von der Gestalt

$$\Phi' = S^{-1}\,\Phi\,S \tag{66}$$

unterworfen ist, wobei nach (58) S zeitunabhängig ist.

Ferner gilt, wie leicht einzusehen [1]), Formel (45) auch für infinitesimale Transformationen $\overline{\mathfrak{N}}$ der Gruppe, d. h. es ist,

---

1) Ersetzt man in

$$\Phi' = \Phi + \frac{1}{\omega}\,[\overline{\mathfrak{N}}, \Phi]$$

Φ durch

$$\widetilde{\Phi} = \Phi + \frac{1}{\omega}\,[\overline{\mathfrak{M}}, \Phi]$$

und Φ' durch

$$\widetilde{\Phi}' = \Phi' + \frac{1}{\omega}\,[\overline{\mathfrak{M}}, \Phi'],$$

so kommt nach leichter Rechnung

$$\widetilde{\Phi}' = \widetilde{\Phi} + \frac{1}{\omega}\left[\overline{\mathfrak{N}} + \frac{1}{\omega}\,[\overline{\mathfrak{M}}, \overline{\mathfrak{N}}],\ \widetilde{\Phi}\right].$$

Vgl. auch E. Noether, Gött. Nachr. 1918, S. 252.

wenn man alle Feldgrößen der infinitesimalen Transformation $\overline{\mathfrak{M}}$ unterwirft,

(67) $$\omega\, \delta^* \overline{\mathfrak{N}} = [\overline{\mathfrak{M}}, \overline{\mathfrak{N}}],$$

daher allgemeiner

(67′) $$\overline{\mathfrak{N}}' = S^{-1} \overline{\mathfrak{N}} S .$$

Aus (66) folgt unmittelbar die Invarianz der kanonischen V.-R. (37). Nach (35) besteht die Hamiltonfunktion aus einem nur von Q und $\mathfrak{P}$ abhängigen Funktional $\mathfrak{H}_0$ und einem Anteil $\overline{\lambda^r \mathfrak{F}_r}$, der nach § 6 eine spezielle infinitesimale Transformation $\overline{\mathfrak{N}}$ darstellt. Wegen (66) und (67′) erleiden also auch die kanonischen Feldgleichungen (38) eine (zeitlich konstante) unitäre Transformation, bei der sie bekanntlich invariant bleiben.

Es bleibt noch übrig, die Variation der linken Seiten $\mathfrak{F}_r$ der Identitäten (29) zu untersuchen. Nach (67) beträgt sie

(68) $$\omega\, \delta^* \mathfrak{F}_r = [\overline{\mathfrak{M}}, \mathfrak{F}_r] .$$

Nun gilt infolgedessen, daß die durch (65) definierte Gruppe $\overline{\mathfrak{S}}$ eine invariante Untergruppe ist,

(68′) $$[\overline{\mathfrak{M}}, \mathfrak{F}_r] = \sum_{i=k+1}^{i=j} \alpha_i^{rs} \frac{d^{j-i} \mathfrak{F}_s}{(dx^4)^{j-i}} .$$

Gemäß (68) und (68′) sind also die $\delta^* \mathfrak{F}_r = 0$, d. h. die eigentlichen Identitäten $\mathfrak{F}_r = 0$ invariant, und zwar vermöge der Identitäten selber und eventuell deren zeitlichen Ableitungen.

§ 9. Erweiterung der Theorie auf den „zweiten Fall“ des § 1

Wir deuten kurz an, wie sich die vorige Theorie auf den am Ende des § 1 definierten „zweiten Fall“ ausdehnt.

Unsere Gruppe habe also die einfache Form:

(69) $$\begin{cases} \delta x^\nu = 0 \\ \delta Q_\alpha = c_{\alpha r}\, \xi^r . \end{cases}$$

Mit

(14) $$\mathfrak{L}' \equiv \frac{d}{d x^\nu} \left[\underline{\mathfrak{f}^{\nu, \alpha \varrho}(Q)\, Q_{\alpha,\, \varrho}} \right]$$

ist nach (12)

(70) $$\delta (\mathfrak{L} + \mathfrak{L}') = 0 .$$

1. Berechnen wir zunächst $\delta\mathfrak{L}'$. Ich behaupte, daß $\delta\mathfrak{L}'$ die Form

$$\delta\mathfrak{L}' = \frac{d}{d x^\nu}\left(\mathfrak{R}^{\alpha\nu}\,\delta Q_\alpha\right) \tag{71}$$

oder

$$\delta\mathfrak{L}' = \frac{d}{d x^\nu}\left(\mathfrak{I}_r{}^\nu\,\xi^r\right) \tag{72}$$

annimmt.

Denn wir bekommen zunächst

$$\delta\mathfrak{L}' = \frac{d}{d x^\nu}\left\{\frac{\partial\,\mathfrak{f}^{\nu,\alpha\varrho}}{\partial Q_\beta}\,c_{\beta r}\,\xi^r\,Q_{\alpha,\varrho} + \mathfrak{f}^{\nu,\alpha\varrho}\,\frac{d\,(c_{\alpha r}\,\xi^r)}{d x^\varrho}\right\};$$

setzen wir

$$\mathfrak{r}^{\alpha\nu} = -\frac{d\,\mathfrak{f}^{\nu,\alpha\varrho}}{d x^\varrho} + Q_{\beta,\varrho}\,\frac{\partial\,\mathfrak{f}^{\nu,\beta\varrho}}{\partial Q_\alpha} \tag{73}$$

und

$$\mathfrak{I}_r{}^\nu = \mathfrak{r}^{\alpha\nu}\,c_{\alpha r}, \tag{74}$$

so wird

$$\delta\mathfrak{L}' = \frac{d}{d x^\nu}\left\{\mathfrak{I}_r{}^\nu\,\xi^r + \frac{d}{d x^\varrho}\left(\mathfrak{f}^{\nu,\alpha\varrho}\,c_{\alpha r}\,\xi^r\right)\right\}. \tag{75}$$

Jetzt benutzen wir (70) und drücken aus, daß die Koeffizienten der zweiten Ableitungen der ξ^r identisch verschwinden. Da $\mathfrak{L}$ keine zweiten Ableitungen der ξ^r enthält, so wird nach (75)

$$\left(\mathfrak{f}^{\nu,\alpha\varrho} + \mathfrak{f}^{\varrho,\alpha\nu}\right)c_{\alpha r} \equiv 0\,. \tag{76}$$

Infolgedessen reduziert sich (75) zu

$$\delta\mathfrak{L}' = \frac{d}{d x^\nu}\left(\mathfrak{I}_r{}^\nu\,\xi^r\right).$$

Setzen wir noch

$$\mathfrak{R}^{\alpha\nu} = \mathfrak{r}^{\alpha\nu} \tag{77}$$

und bemerken, daß statt (74) auch

$$\mathfrak{I}_r{}^\nu = \mathfrak{R}^{\alpha\nu}\,c_{\alpha r} \tag{74}$$

geschrieben werden darf, so haben wir die Formeln (71) und (72) bewiesen.

2. Jetzt stellen wir die Analoga der Identitäten (28) auf, die im ersten Falle auch eigentliche Identitäten (29) enthielten.

Dazu haben wir bloß die Koeffizienten der $\dfrac{d\,\xi^r}{d\,x^\nu}$ in (70) gleich Null zu setzen. Das liefert uns

$$(\mathfrak{P}^{\alpha\nu} + \mathfrak{R}^{\alpha\nu})\, c_{\alpha r} = 0\,, \tag{78}$$

Insbesondere für $\nu = 4$:

$$(\mathfrak{P}^{\alpha} + \mathfrak{R}^{\alpha 4})\, c_{\alpha r} = 0\,,$$

oder, indem wir wiederum

$$\mathfrak{F}_r \equiv \mathfrak{P}^{\alpha}\, c_{\alpha r} \quad \text{und} \quad \mathfrak{J}_r^{\,4} \equiv \mathfrak{J}_r$$

setzen,

$$\mathfrak{F}_r + \mathfrak{J}_r = 0\,. \tag{79}$$

3. Die Identitäten (79) sind eigentliche, d. h. es ist

$$\frac{\partial\, \mathfrak{J}_r}{\partial\, \dot{Q}_\alpha} = 0\,. \tag{80}$$

Allgemeiner wollen wir statt (80) beweisen, daß

$$\frac{\partial\, (\mathfrak{r}^{\beta\varrho}\, c_{\beta r})}{\partial\, Q_{\alpha,\nu}} = - \frac{\partial\, (\mathfrak{r}^{\beta\nu}\, c_{\beta r})}{\partial\, Q_{\alpha,\varrho}}\,, \tag{80'}$$

woraus (80) nach (74) für $\nu = \varrho = 4$ folgt.

Zu dem Zweck setzen wir den Koeffizienten von ξ_r in (70) gleich Null: es ist ein in den zweiten Ableitungen $Q_{\alpha,\varrho\nu}$ linearer Ausdruck, mit nur von Q abhängigen Koeffizienten. Da dieser Ausdruck für beliebige $Q_{\alpha,\varrho\nu}$ identisch verschwindet, können wir insbesondere den $Q_{\alpha,\varrho\nu}$ c-Zahlwerte zuschreiben und dann die Koeffizienten der $Q_{\alpha,\varrho\nu}$ getrennt gleich Null setzen. Unter Benutzung der für beliebiges $\mathfrak{K}^\varrho(Q_\alpha;\, Q_{\alpha,\nu})$ gültigen Formel:

$$\frac{\partial}{\partial\, Q_{\alpha,\nu}}\, \frac{d}{d\,x^\varrho}\, \mathfrak{K}^\varrho(Q_\alpha;\, Q_{\alpha,\nu}) = \frac{\partial\, \mathfrak{K}^\nu}{\partial\, Q_\alpha} + \frac{d}{d\,x^\varrho}\, \frac{\partial\, \mathfrak{K}^\varrho}{\partial\, Q_{\alpha,\nu}}\,, \tag{81}$$

finden wir für diese Koeffizienten nach (71) und (73)

$$\frac{\partial\, (\mathfrak{r}^{\beta\varrho}\, c_{\beta r})}{\partial\, Q_{\alpha,\nu}} + \frac{\partial\, (\mathfrak{r}^{\beta\nu}\, c_{\beta r})}{\partial\, Q_{\alpha,\varrho}}\,,$$

deren Nullsetzen (80') ergibt.

Gemäß (81) folgt übrigens aus (73)

$$\frac{\partial\, \mathfrak{r}^{\alpha\nu}}{\partial\, Q_{\beta,\varrho}} = \frac{\partial\, \mathfrak{f}^{\nu,\beta\varrho}}{\partial\, Q_\alpha} - \frac{\partial\, \mathfrak{f}^{\nu,\alpha\varrho}}{\partial\, Q_\beta} = \frac{\partial\, \mathfrak{R}^{\alpha\nu}}{\partial\, Q_{\beta,\varrho}}\,; \tag{82}$$

danach können wir statt (80′) auch

(83)
$$\frac{\partial(\mathfrak{R}^{\beta\varrho}c_{\beta r})}{\partial Q_{\alpha,\nu}} = -\frac{\partial(\mathfrak{R}^{\beta\nu}c_{\beta r})}{\partial Q_{\alpha,\varrho}}$$

schreiben.

4. Die Ausführungen der §§ 3 und 4 sind auf den jetzigen Fall wörtlich übertragbar, *indem* $\mathfrak{P}^{\alpha} + \mathfrak{R}^{\alpha 4}$ *die Rolle von* $\mathfrak{P}^{\alpha}$ *vertritt.*

Der im § 5 abgeleitete Ausdruck $\overline{\mathfrak{M}}$ der infinitesimalen Transformation erleidet eine analoge Modifikation, da $\mathfrak{P}^{\alpha\nu}$ jetzt keine Tensordichte mehr ist.[1])

Vielmehr ist jetzt nach (50) und (70)

$$\delta\mathfrak{P}^{\alpha\nu} = -\mathfrak{P}^{\beta\nu}\frac{\partial\,\delta Q_{\beta}}{\partial Q_{\alpha}} - \frac{\partial(\delta\mathfrak{L}')}{\partial Q_{\alpha,\nu}};$$

nach (71) ist aber, gemäß (81),

$$\frac{\partial(\delta\mathfrak{L}')}{\partial Q_{\alpha,\nu}} = \frac{\partial(\mathfrak{R}^{\beta\nu}\,\delta Q_{\beta})}{\partial Q_{\alpha}} + \frac{d}{dx^{\varrho}}\frac{\partial(\mathfrak{R}^{\beta\varrho}\,\delta Q_{\beta})}{\partial Q_{\alpha,\nu}},$$

d. h. mit Rücksicht auf (83)

(84)
$$\delta\mathfrak{P}^{\alpha\nu} = -\mathfrak{P}^{\beta\nu}\frac{\partial\,\delta Q_{\beta}}{\partial Q_{\alpha}} - \frac{\partial(\mathfrak{R}^{\beta\nu}\,\delta Q_{\beta})}{\partial Q_{\alpha}} + \frac{d}{dx^{\varrho}}\frac{\partial(\mathfrak{R}^{\beta\nu}\,\delta Q_{\beta})}{\partial Q_{\alpha,\varrho}}.$$

Insbesondere wird das wegen (80) für $\nu = 4$

$$\delta\mathfrak{P}^{\alpha} = -\mathfrak{P}^{\beta}\frac{\partial\,\delta Q_{\beta}}{\partial Q_{\alpha}} - \frac{\partial(\mathfrak{R}^{\beta 4}\,\delta Q_{\beta})}{\partial Q_{\alpha}} + \frac{d}{dx^{\varrho}}\frac{\partial(\mathfrak{R}^{\beta 4}\,\delta Q_{\beta})}{\partial Q_{\alpha,\varrho}},$$

d. h.

(85)
$$\omega\,\delta\mathfrak{P}^{\alpha} = [\overline{\mathfrak{N}}, \mathfrak{P}^{\alpha}],$$

mit

(86)
$$\overline{\mathfrak{N}} = (\mathfrak{P}^{\alpha} + \mathfrak{R}^{\alpha 4})\,\delta Q_{\alpha}.$$

Wiederum wegen (80) ist auch

(87)
$$\omega\,\delta Q_{\alpha} = [\overline{\mathfrak{N}}, Q_{\alpha}],$$

so daß wir in $\overline{\mathfrak{N}}$ die gesuchte Erweiterung von $\overline{\mathfrak{M}}$ haben.

---

1) Obwohl weder $\mathfrak{P}^{\alpha\nu}$ noch $\mathfrak{R}^{\alpha\nu}$ Tensordichten sind, läßt sich leicht zeigen, daß $\mathfrak{P}^{\alpha\nu} + \mathfrak{R}^{\alpha\nu}$ dennoch eine Tensordichte ist.

Aus dem Ausdruck (86) folgt genau wie im § 6, daß die linken Seiten $\mathfrak{F}_r + \mathfrak{I}_r$ der eigentlichen Identitäten auf Grund derselben untereinander vertauschbar sind.

Die Überlegungen von § 7 über die zeitliche Konstanz von $\overline{\mathfrak{M}}$, sowie der Kovarianzbeweis von § 8, lassen sich ohne weiteres auf $\overline{\mathfrak{N}}$ übertragen. Insbesondere spielen hier, da $j = 0$ vorausgesetzt wurde, die Identitäten $\mathfrak{F}_r + \mathfrak{I}_r = 0$ die Rolle der fehlenden Feldgleichungen.

§ 10. Bemerkung über die gleichzeitige Behandlung mehrerer Gruppen

Auf den Fall, daß die Lagrangefunktion mehrere Gruppen gestattet, ist die obige Theorie ohne weiteres anwendbar, wenn man bedenkt, daß die infinitesimale Transformation des direkten Produktes aller betrachteten Gruppen sich additiv aus denen der einzelnen Gruppen zusammensetzt. Insbesondere sind die sich auf jede einzelne Gruppe beziehenden $\mathfrak{F}_r$ nicht nur untereinander (vermöge $\mathfrak{F}_r = 0$) vertauschbar, sondern auch mit den zu den anderen Einzelgruppen gehörigen $\mathfrak{F}_r$. Es ist auch zulässig, daß der „erste Fall" ($\mathfrak{L}$ ist eine Dichte) für gewisse Einzelgruppen, der im § 9 behandelte „zweite Fall" dagegen für gewisse anderen auftritt. Dann sind für die letzteren die $\mathfrak{F}_r$ einfach durch $\mathfrak{F}_r + \mathfrak{I}_r$ zu ersetzen: sie sind wiederum nicht nur untereinander, sondern auch mit den anderen $\mathfrak{F}_r$ vertauschbar.

Diese Bemerkung hat zur Folge, daß man die einzelnen Gruppen, welche eine gegebene Lagrangefunktion gestattet, getrennt behandeln kann.

Zweiter Teil: **Anwendungen**

§ 11. Die Lagrangefunktion

Wir wollen eine Lagrangefunktion aufstellen, die sowohl das elektromagnetische und Materiefeld als auch das Gravitationsfeld umfaßt. Was das letztere betrifft, so übernehmen wir die von Fock[1]) und Weyl[2]) vorgeschlagene Theorie des Einkörperproblems: das Gravitationsfeld beschreiben wir durch

1) V. Fock, Ztschr. f. Phys. **57**. S. 261. 1929.

2) H. Weyl, Ztschr. f. Phys. **56**. S. 330. 1929.

Angabe in jedem Punkt von vier orthogonalen Vektoren $h_{i,\nu}$ $(i = 1, 2, 3, 4)$ und wir fordern, daß die Naturgesetze kovariant sind gegenüber einer *vom Punkte abhängigen* Lorentztransformation der „Vierbeine" $h_{i,\nu}$; diese Kovarianz, die wir nach Levi-Cività [1]) „echte Beinkovarianz" nennen, unterscheidet sich wesentlich von der in der Einsteinschen Theorie des Fernparallelismus geforderten „lokalen Beinkovarianz", nach der alle Vierbeine miteinander starr verbunden sind (*konstante* Lorentztransformation der Vierbeine). Im Einklang mit Fock (und im Gegensatz zu Weyl) beschreiben wir das Materiefeld durch *vier*komponentige Wellenfunktionen $\psi \equiv (\psi_1, \psi_2, \psi_3, \psi_4)$. Für das elektromagnetische Feld wählen wir als Variable die Komponenten φ_μ des Viererpotentials.[2])

Die Lagrangefunktion setzt sich additiv zusammen aus drei Anteilen, die den drei genannten Feldern entsprechen (und gleichzeitig die Wechselwirkungen der Felder aufeinander enthalten).

Wenn

$$E_{\mu\nu} = \frac{\partial \varphi_\nu}{\partial x^\mu} - \frac{\partial \varphi_\mu}{\partial x^\nu} \tag{88}$$

den elektromagnetischen Feldtensor darstellt, so ist der Strahlungsanteil der Lagrangefunktion

$$\mathfrak{E} = \tfrac{1}{4} E_{\mu\nu}\, \mathfrak{E}^{\mu\nu}; \tag{89}$$

dabei bedeutet

$$\mathfrak{E}^{\mu\nu} = E^{\mu\nu} h'$$

wo h' die Determinante der $h_{i,\nu}$ und $E^{\mu\nu}$ die kontravarianten Komponenten des Tensors $E_{\mu\nu}$ bezeichnet.

Um den Materieanteil aufzuschreiben, legen wir ein spezielles System Diracscher Matrizen fest.[3]) Gehen wir aus von den Paulischen Matrizen

$$\varrho_1 = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \qquad \varrho_2 = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}, \qquad \varrho_3 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \tag{90}$$

1) Berliner Berichte 1929, S. 137.

2) Da wir $x^4 = ct$ gesetzt haben, ist $\varphi_4 = -\varphi$, wo φ das skalare Potential darstellt.

3) Dasselbe weicht von dem Fockschen (a. a. O.) nur unwesentlich ab. Der wesentliche Zug der Spezialisierung ist $\alpha_4 = 1$ (in der Fockschen Bezeichnung $\alpha_0 = 1$).

so setzen wir

$$\left\{ \begin{aligned} \alpha_{\bar{l}} &= -i \begin{pmatrix} \varrho_{\bar{l}} & 0 \\ 0 & -\varrho_{\bar{l}} \end{pmatrix} \quad (\bar{l} = 1, 2, 3) \\ \alpha_4 &= 1 . \end{aligned} \right. \tag{91}$$

Führen wir noch die Bezeichnung

$$e_{\bar{k}} = -1, \quad e_4 = 1 \tag{92}$$

ein, so sind die Matrizen α_l hermitisch und haben die Vertauschungseigenschaft

$$\alpha_m \alpha_k e_k + \alpha_k \alpha_m e_m = 2 e_m \delta_{mk} . \tag{93}$$

Ferner brauchen wir noch die Matrix

$$\sigma = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{94}$$

(worin die Einser zweireihige Einheitsmatrizen darstellen).

In bezug auf die lateinischen Indizes ist über zweimal auftretende Indizes zu summieren, wobei die Faktoren e_k zur Abzählung der Indizes unberücksichtigt bleiben sollen. Neben die $h_{i,\nu}$ treten die kontravarianten h_i^{ν} und es gelten die Relationen

$$\left\{ \begin{aligned} h_k^{\nu} h_{l,\nu} &= e_k \delta_{kl} , \\ e_k h_k^{\nu} h_{k,\mu} &= \delta_{\mu} , \end{aligned} \right. \tag{95}$$

welche die Orthogonalität der Vierbeine im Raume mit der Maßbestimmung

$$g_{\mu\nu} = e_k h_{k,\mu} h_{k,\nu} \tag{96}$$

ausdrücken.

Bedeutet noch

$$\eta^{l}_{\varrho\sigma} = \frac{\partial h_{l,\varrho}}{\partial x^{\sigma}} - \frac{\partial h_{l,\sigma}}{\partial x^{\varrho}} \tag{97}$$

und

$$2 \gamma_{mkl} = \left(\eta^{l}_{\varrho\sigma} h^{\sigma}_{m} h^{\varrho}_{k} + \eta^{m}_{\varrho\sigma} h^{\sigma}_{l} h^{\varrho}_{k} + \eta^{k}_{\varrho\sigma} h^{\sigma}_{m} h^{\varrho}_{l} \right) h' , \tag{98}$$

ferner

$$C_l = \frac{1}{4} e_k \alpha_m \alpha_k \gamma_{mkl} + \frac{e}{\omega} \varphi_{\sigma} h_l^{\sigma} h' , \quad \left(\omega = \frac{hc}{2\pi i} \right) \tag{99}$$

$$\gamma^{\sigma} = e_k \alpha_k h^{\sigma}_{k} h' , \tag{100}$$

so lautet der Materieanteil der Lagrangefunktion:

$$\boldsymbol{R} \omega \psi^* \left(\gamma^{\sigma} \frac{\partial \psi}{\partial x^{\sigma}} - e_l \alpha_l C_l \psi \right) - m c^2 \psi^* \sigma \psi h' . \tag{101}$$

[x^* = komplex konjugiertes von x, $\boldsymbol{R}\,x$ = Realteil von x, $\boldsymbol{I}\,x$ = Imaginärteil von x].

Nun ist [vgl. Fock, a. a. O. Formel (24)]

$$e_l\,(\alpha_l\,C_l + C_l^\dagger \alpha_l) = -\frac{\partial \gamma^\sigma}{\partial x^\sigma}$$

und infolgedessen

$$\boldsymbol{I}\,\omega\,\psi^* \left(\gamma^\sigma \frac{\partial \psi}{\partial x^\sigma} - e_l\,\alpha_l\,C_l\,\psi\right) = \frac{\omega}{2}\,\frac{\partial}{\partial x^\sigma}(\psi^*\,\gamma^\sigma\,\psi). \tag{102}$$

Wir können also für den Materieanteil statt (101)

$$\mathfrak{W} = \omega\,\psi^* \left(\gamma^\sigma \frac{\partial \psi}{\partial x^\sigma} - e_l\,\alpha_l\,C_l\,\psi\right) - m\,c^2\,\psi^*\,\sigma\,\psi\,h' \tag{103}$$

setzen.

Für den Gravitationsanteil nehmen wir $\frac{1}{2\varkappa}\,\mathfrak{G}$, wo $\varkappa = \frac{8\pi f}{c^4}$ (f = Newtonsche Gravitationskonstante) und

$$\left\{\begin{aligned} \mathfrak{G} = {} & e_k e_l\,\eta^l_{\varrho\sigma}\,h_l{}^\varrho\,h_k{}^{\varrho'}\,g^{\sigma\sigma'}\,h'\,\eta^k_{\varrho'\sigma'} - \tfrac{1}{2}\,e_k\,e_l\,\eta^l_{\varrho\sigma}\,h_l{}^{\varrho'}\,h_k{}^\varrho\,g^{\sigma\sigma'}\,h'\,\eta^k_{\varrho'\sigma'} \\ & - \tfrac{1}{4}\,e_l\,\eta^l_{\varrho\sigma}\,g^{\sigma\sigma'}\,g^{\varrho\varrho'}\,h'\,\eta^l_{\varrho'\sigma'}; \end{aligned}\right. \tag{104}$$

wie man leicht nachrechnet (vgl. etwa Weyl, a. a. O.) unterscheidet sich $\mathfrak{G}$ von der skalaren Krümmungsdichte $\mathfrak{R}$ um eine Divergenz:

$$\mathfrak{R} = \mathfrak{G} - 2\,\frac{d}{dx^\nu}\underline{\left(e_l\,h_l{}^\nu\,\frac{\partial\,(h_l{}^\sigma h')}{\partial x^\sigma}\right)}. \tag{105}$$

Insgesamt ist also

$$\mathfrak{L} = \frac{1}{2\varkappa}\,\mathfrak{G} + \mathfrak{E} + \mathfrak{W}. \tag{106}$$

Zum Unterschied gegenüber der gewöhnlichen Form der Relativitätstheorie, wo die Feldgrößen $g_{\mu\nu}$ keine Vektoren, sondern Tensoren 2. Stufe waren, sind in (105) die beiden Bestandteile $\mathfrak{G}$ und $2\,\frac{d}{dx^\nu}\underline{\left(e_l\,h_l{}^\nu\,\frac{\partial\,(h_l{}^\sigma h')}{\partial x^\sigma}\right)}$ von $\mathfrak{R}$ skalare Dichten bezüglich der allgemeinen relativistischen Transformationsgruppe. Dagegen ist nicht $\mathfrak{G}$ allein, sondern erst $\mathfrak{R}$ echt beininvariant.

§ 12. Die Eichinvarianzgruppe

Die einfachste Gruppe, die unsere Funktion $\mathfrak{L}$ gestattet, ist die Eichinvarianzgruppe, bei welcher die x^ν und die Variabeln $h_{i,\nu}$ invariant bleiben, während sich die φ_ν und ψ folgendermaßen transformieren:

$$(107)\qquad \begin{cases} \delta\varphi_\nu = \dfrac{\partial \xi}{\partial x^\nu}, \\ \delta\psi = -\dfrac{e}{\omega}\,\xi\,\psi. \end{cases}$$

Gegenüber dieser Gruppe ist $\delta\mathfrak{L} = 0$.

Setzen wir, um den Vergleich mit der allgemeinen Theorie zu erleichtern[1]),

$$\varphi_\nu = Q_\nu, \qquad \psi = Q_5,$$

so haben wir

$$c_\nu{}^\mu = \delta_\nu{}^\mu, \qquad c_5{}^\mu = 0$$

und folglich als einzige eigentliche Identität

$$(108)\qquad \mathfrak{P}^4 = 0.$$

Das folgt natürlich aus der direkten Ausrechnung der $\mathfrak{P}^{\alpha\nu}$:

$$(109)\qquad \begin{cases} \mathfrak{P}^{\mu\nu} = \mathfrak{F}^{\nu\mu}, \\ \mathfrak{P}^{5\nu} = \omega\,\psi^*\gamma^\nu. \end{cases}$$

Um dieses einfache Beispiel weiter zu diskutieren, sehen wir zunächst von der Gravitation ab, d. h. setzen wir $h_{i,\nu} = \delta_{i\nu}$.

Die Hamiltonfunktion hat dann die Form

$$(110)\qquad \mathfrak{H} = \mathfrak{H}_0 + \lambda\,\mathfrak{P}^4,$$

wo $\mathfrak{H}_0$ z. B. die in H. P. II gewählte spezielle Hamiltonfunktion ist, welche $\mathfrak{P}^4$ nicht enthält.

Die Feldgleichungen lauten

$$(111)\qquad \begin{cases} \omega\,\dot{Q}_{\bar\nu} = [\overline{\mathfrak{H}}_0, Q_{\bar\nu}], \\ \dot{Q}_4 = \lambda, \\ \omega\,\dot{Q}_5 = [\overline{\mathfrak{H}}_0, Q_5]; \end{cases}$$

$$(112)\qquad \omega\,\dot{\mathfrak{P}}^\alpha = [\overline{\mathfrak{H}}_0, \mathfrak{P}^\alpha], \quad (\alpha = 1, \ldots, 5);$$

---

1) Da die ψ nicht hermitisch sind, so sind dem allgemeinen Schema geringe Modifikationen anzubringen, um es auch diesen Variablen anzupassen. Darauf brauchen wir aber nicht näher einzugehen.

da ferner $j = 1$ ist, so haben wir als Nebenbedingung [1] außer (108)

$$[\overline{\mathfrak{H}}_0, \mathfrak{P}^4] = 0: \tag{113}$$

Also bleibt in (111) λ wesentlich unbestimmt und die vierte Gleichung (112) wird durch (113) ersetzt.

Die infinitesimale Transformation $\overline{\mathfrak{M}}$ lautet hier:

$$\overline{\mathfrak{M}} = \int \left\{ \frac{\partial \xi}{\partial x^\nu} \mathfrak{E}^{4\nu} - e\, \xi\, \psi^* \gamma^4 \psi \right\} d x^1 d x^2 d x^3$$

oder durch partielle Integration

$$\overline{\mathfrak{M}} = \int \left\{ \frac{\partial \xi}{\partial x^4} \mathfrak{P}^4 - \xi \left[\frac{\partial \mathfrak{E}^{4\nu}}{\partial x^\nu} + e\, \psi^* \gamma^4 \psi \right] \right\} d x^1 d x^2 d x^3.$$

Die eckige Klammer ist nichts anderes als $1/\omega\, [\overline{\mathfrak{H}}_0, \mathfrak{P}^4]$ oder $\dot{\mathfrak{P}}^4$, so daß

$$\overline{\mathfrak{M}} = \int \left\{ \frac{\partial \xi}{\partial x^4} \mathfrak{P}^4 - \xi \frac{d\, \mathfrak{P}^4}{d x^4} \right\} d x^1 d x^2 d x^3, \tag{114}$$

in Übereinstimmung mit (63').

Nach der allgemeinen Theorie muß $\dot{\mathfrak{P}}^4 = 0$ vermöge der Feldgleichungen und Identitäten identisch erfüllt sein; das ist in der Tat die Kontinuitätsgleichung der Elektrizität.

§ 13. Die allgemeine relativistische Kovarianz

Bei einer beliebigen Koordinatentransformation

$$\delta x^\nu = \xi^\nu \tag{115}$$

ist

$$\delta h_{i,\nu} = - h_{i,\mu} \frac{\partial \xi^\mu}{\partial x^\nu}, \tag{115'}$$

ferner

$$\left\{ \begin{aligned} \delta \varphi_\nu &= - \varphi_\mu \frac{\partial \xi^\mu}{\partial x^\nu}, \\ \delta \psi &= 0. \end{aligned} \right. \tag{115''}$$

Die Lagrangefunktion verhält sich dabei wie eine skalare Dichte.

Behalten wir die Bezeichnungen des vorigen Paragraphen für die zu den φ_ν und ψ konjugierten Impulsen bei, und

[1] Mit den Bezeichnungen von H.P.II lautet (113) $C = 0$.

stellen wir die zu den $h_{i,\nu}$ gehörigen $\mathfrak{P}^{\alpha\mu}$ durch $\mathfrak{P}_i^{\nu\mu}$ dar, so lauten die uneigentlichen Identitäten (28) im jetzigen Falle:

$$\varphi_\varrho(\mathfrak{E}^{\mu\nu} + \mathfrak{E}^{\nu\mu}) + h_{i,\varrho}(\mathfrak{P}_i^{\nu\mu} + \mathfrak{P}_i^{\mu\nu}) = 0,$$

mit Rücksicht darauf, daß $\mathfrak{E}^{\mu\nu} + \mathfrak{E}^{\nu\mu} = 0$ ist, reduzieren sie sich zu

$$\mathfrak{P}_i^{\nu\mu} + \mathfrak{P}_i^{\mu\nu} = 0 \tag{116}$$

und haben also die vier eigentlichen Identitäten

$$\mathfrak{P}_i^4 = 0 \tag{117}$$

zur Folge.

Die direkte Berechnung ergibt in der Tat (116), da $\mathfrak{G}$ und $\mathfrak{W}$ von den $h_{i,\nu,\mu}$ nur durch die $\eta^i_{\tau\sigma}$ abhängen und

$$\frac{\partial \eta^i_{\varrho\sigma}}{\partial h_{i,\nu,\mu}} = \delta^\nu_\varrho \delta^\mu_\sigma - \delta^\mu_\varrho \delta^\nu_\sigma,$$

d. h. antisymmetrisch in μ und ν ist. Man findet

$$\left\{\begin{aligned} \mathfrak{P}_i^{\nu\mu} = \Big[& \eta^i_{\varrho\sigma} g^{\nu\sigma} g^{\mu\varrho} + 2 e_l \eta^l_{\varrho\sigma} h^\varrho_l \left(g^{\sigma\mu} h_i^\nu - g^{\sigma\nu} h_i^\mu\right) \\ & - e_l \eta^l_{\varrho\sigma} h_i^\varrho (h_l^\nu g^{\sigma\mu} - h_l^\mu g^{\sigma\nu})\Big] e_i h' \cdot \frac{1}{2\varkappa} \\ & - \boldsymbol{R} \frac{\omega}{4} \psi^* \alpha_l \alpha_m \alpha_k \psi e_l e_k \frac{\partial \gamma_{mkl}}{\partial h_{i,\nu,\mu}}. \end{aligned}\right. \tag{118}$$

Die infinitesimale Transformation $\overline{\mathfrak{M}}$ nimmt die Form an:

$$\overline{\mathfrak{M}} = -\int d x^1 d x^2 d x^3 \left\{ \frac{\partial \xi^\mu}{\partial x^4} (h_{i,\mu} \mathfrak{P}_i^4 + \varphi_\mu \mathfrak{P}^4) \right.$$
$$\left. - \xi^\mu \left[\frac{\partial}{\partial x^{\bar\nu}} (h_{i,\mu} \mathfrak{P}_i \ + \varphi_\mu \mathfrak{P}^{\bar\nu}) + \mathfrak{G}_\mu \right] \right\};$$

betrachten wir insbesondere die Translation $\xi^\mu = \varepsilon^\mu = \mathrm{const}$ so liefert $\frac{d\overline{\mathfrak{M}}}{dx^4} = 0$ den Energieimpulssatz

$$\overline{\mathfrak{G}}_\mu = \mathrm{const};$$

für eine lineare Transformation, bekommen wir daraus eine Verallgemeinerung der Drehimpulssätze (vgl. H. P. II, S. 177).

Nach der allgemeinen Theorie (§ 7) liefert das Nullsetzen des Koeffizienten von ξ^μ eine Nebenbedingung:

$$\mathfrak{G}_\mu + \frac{\partial}{\partial x^{\bar\nu}} (h_{i,\mu} \mathfrak{P}_i^{\bar\nu} + \varphi_\mu \mathfrak{P}^{\bar\nu}) = 0. \tag{119}$$

F. Klein[1]) hat bereits in anderem Zusammenhang bemerkt, daß (119) mit 4 Feldgleichungen äquivalent ist.

§ 14. Die echte Beinkovarianz

Für diese Gruppe ist

$$(120)\qquad \begin{cases} \delta x^\nu = 0\,, & \delta \varphi_\mu = 0\,, \\ \delta h_{i,\nu} = e_k \xi_{ik} h_{k,\nu}\,, & (\xi_{ik} = -\xi_{ki}) \end{cases}$$

und, wie man auf Grund von (120) leicht findet,

$$(120')\qquad \delta \psi = \tfrac{1}{4} e_k \xi_{ik} \alpha_i \alpha_k \psi\,.$$

Hier haben wir ein Beispiel des im § 9 behandelten „zweiten Falles" vor uns. Denn $\mathfrak{G}$ ist nur lokalbeininvariant und erst $\mathfrak{R}$ ist echt beininvariant; $\mathfrak{W}$ und $\mathfrak{E}$ sind echt beininvariant.

Um nach Formel (74) $\mathfrak{J}_r^{\,\nu} \equiv \mathfrak{J}^\nu_{(ik)}$ auszurechnen, ist es am zweckmäßigsten, vorübergehend als Variable

$$Q_l^{\,\alpha} \equiv h' h_l^{\,\alpha}$$

zu wählen. Dann ist nach (105)

$$\mathfrak{f}_l^{\,\nu,\,\alpha\varrho} = -\frac{1}{\varkappa}\, e_l h_l^{\,\nu} \delta_\varrho^{\,\alpha}$$

und nach (120)

$$c_{l,\,\alpha(ik)} = \delta_{il} e_k h_k^{\,\alpha} h' - \delta_{kl} e_i h_i^{\,\alpha} h'.$$

Unter Berücksichtigung von (71) findet man daraus leicht

$$(121)\qquad \mathfrak{J}^\nu_{(ik)} = \frac{1}{\varkappa}\, e_i e_k \frac{d}{dx^\varrho} \left(h' h_i^{\,\nu} h_k^{\,\varrho} - h' h_i^{\,\varrho} h_k^{\,\nu}\right).$$

Zur Berechnung von $\mathfrak{F}_r \equiv \mathfrak{F}_{(ik)}$ ist es bequemer, zu den ursprünglichen Variablen $Q_{l,\alpha} \equiv h_{l,\alpha}$ und $Q_5 = \psi$ zurückzukehren. Dann ist

$$c_{l,\,\alpha(ik)} = \delta_{il} e_k h_{k,\alpha} - \delta_{kl} e_i h_{i,\alpha}\,,$$

ferner nach (120′)

$$c_{5\,(ik)} = \tfrac{1}{4}\left(e_k \alpha_i \alpha_k \psi - e_i \alpha_k \alpha_i \psi\right)$$

zu setzen. Danach ist

$$\mathfrak{F}_{(ik)} = \underline{\mathfrak{P}_i^{\,\nu} e_k h_{k,\nu} - \mathfrak{P}_k^{\,\nu} e_i h_{i,\nu}} + \frac{\omega}{4} e_k e_l h' h_l^{\,4} \psi^* (\alpha_l \alpha_i \alpha_k - \alpha_l \alpha_k \alpha_i) \psi.$$

---

1) Gött. Nachr. 1918, S. 185.

Nun ist nach (98)

$$\frac{\partial \gamma_{m\,j\,l}}{\partial h_{i,\,\nu,\,4}}\, e_k\, h_{k,\,\nu} - \frac{\partial \gamma_{m\,j\,l}}{\partial h_{k,\,\nu,\,4}}\, e_i\, h_{i,\,\nu} = h'\, h_l^{\,4}\,(\delta_{i\,m}\,\delta_{k\,j} - \delta_{i\,j}\,\delta_{k\,m});$$

setzt man also gemäß (118)

$$\mathfrak{P}_i^{\,\nu} = \tilde{\mathfrak{P}}_i^{\,\nu} - \boldsymbol{R}\,\frac{\omega}{4}\, e_l\, \psi^*\, \alpha_l\, \alpha_m\, \alpha_j\, \psi\, e_j\, \frac{\partial \gamma_{m\,j\,l}}{\partial h_{i,\,\nu,\,4}}, \tag{122}$$

wo also $\tilde{\mathfrak{P}}_i^{\,\nu}$ die Impulse bei Abwesenheit von Materie darstellen, so reduzieren sich die $\mathfrak{J}_{(ik)}$ auf

$$\mathfrak{J}_{(ik)} = \underline{\tilde{\mathfrak{P}}_i^{\,\nu}\, e_k\, h_{k,\,\nu} - \tilde{\mathfrak{P}}_k^{\,\nu}\, e_i\, h_{i,\,\nu}}, \tag{123}$$

wie es auch sein muß.

Gemäß (121) und (123) lauten die sechs eigentlichen Identitäten

$$\underline{\tilde{\mathfrak{P}}_i^{\,\nu}\, e_k\, h_{k,\,\nu} - \tilde{\mathfrak{P}}_k^{\,\nu}\, e_i\, h_{i,\,\nu}} + \frac{1}{\varkappa}\, e_i\, e_k\, \frac{d}{d x^\varrho}\,(h' h_i^{\,4} h_k^{\,\varrho} - h' h_i^{\,\varrho} h_k^{\,4}) = 0, \tag{124}$$

die sich auch direkt aus (118) ergeben.

§ 15. Ergänzende Bemerkungen über das Gravitations- und Materiefeld

1. Nachdem wir in den vorigen Paragraphen skizziert haben, wie die Fock-Weylsche Theorie des Einkörperproblems gequantelt werden kann, möchten wir auf einen Punkt dieses Einkörpermodells kurz eingehen, der bei Fock und bei Weyl verschieden behandelt ist, nämlich die Aufstellung des Impulsenergietensors $\mathfrak{T}_i^{\,\nu}$ der Materie. Der Focksche Ansatz, der zu einem unsymmetrischen Tensor führt, scheint uns unzweckmäßig und wir ziehen die Weylsche Definition

$$\mathfrak{T}_i^{\,\nu} = \frac{\delta\,\mathfrak{W}}{\delta\, h_{i,\,\nu}} \tag{125c}$$

vor, da sie auf Grund der Feldgleichungen einen symmetrischen Tensor liefert. Da aber Weyl mit einem zweikomponentigen ψ operiert, während wir mit Fock bei der Vierkomponententheorie bleiben wollen, so wird es wohl nicht überflüssig sein, die Weylsche Berechnung von $\mathfrak{T}_i^{\,\nu}$ hier mutatis mutandis zu wiederholen.

Die Symmetrie von $\mathfrak{T}_i^{\,\nu}$ folgt unmittelbar aus $\delta\,\mathfrak{W} = 0$, wo δ die Variation (120), (120') ist. Denn man bekommt daraus durch Nullsetzen des Koeffizienten von ξ_{ik}

$$\mathfrak{T}_i^{\,\nu} e_k h_{k,\nu} - \mathfrak{T}_k^{\,\nu} e_i h_{i,\nu} = -\frac{1}{2}\,\boldsymbol{R}\,\frac{\delta\,\mathfrak{W}}{\delta\,\psi}\,(e_k\,\alpha_i\,\alpha_k - e_i\,\alpha_k\,\alpha_i)\,\psi,$$

d. h.

$$\mathfrak{T}_i^{\,\nu} e_k h_{k,\nu} - \mathfrak{T}_k^{\,\nu} e_i h_{i,\nu} = 0$$

auf Grund der Feldgleichungen

$$\frac{\delta\,\mathfrak{W}}{\delta\,\psi} = 0 \quad \text{und} \quad \frac{\delta\,\mathfrak{W}}{\delta\,\psi^*} = 0.$$

Diese Gleichung drückt aus, daß der Tensor

$$\mathfrak{T}''_{ik} = e_i\,e_k\,\mathfrak{T}_i^{\,\nu}\,h_{k,\nu}$$

symmetrisch in bezug auf i und k ist.

Statt (125c) können wir ebensogut

$$\text{(126c)} \qquad \mathfrak{T}_i^{\,\nu} = \frac{\delta\,\boldsymbol{R}\,\mathfrak{W}}{\delta\,h_{i,\nu}}$$

setzen, was uns einen reellen Tensor $\mathfrak{T}_i^{\,\nu}$ ergeben wird. Bequemer berechnen wir

$$\text{(127c)} \qquad \mathfrak{T}'_{i,\nu} = \frac{\delta\,\boldsymbol{R}\,\mathfrak{W}}{\delta\,h_i^{\,\nu}} = -\,\mathfrak{T}_k^{\,\varrho}\,e_k\,h_{k,\nu}\,h_{i,\varrho} \equiv e_i\,h'\,T'_{i,\nu}.$$

Auf Grund von (103) finden wir

$$\text{(128c)} \left\{ \begin{aligned} T'_{i,\nu} &= \boldsymbol{R}\,\omega\,\psi^*\,\alpha_i\,\frac{\partial\,\psi}{\partial\,x^\nu} - e\,\psi^*\,\alpha_i\,\psi\,\varphi_\nu - h_{i,\nu} W \\ &\quad + \boldsymbol{R}\,\frac{\omega}{4}\,e_k\,h_k^{\,\varrho}\,h_{m,\nu}\,\frac{\partial}{\partial\,x^\varrho}\,\{\psi^*\,\alpha_i\,\alpha_m\,\alpha_k\,\psi\} \\ &\quad - \boldsymbol{R}\,\frac{\omega}{4}\,e_i e_l e_k\,\psi^*\,\alpha_l\,\alpha_m\,\alpha_k\,\psi \left\{ \frac{\partial\,\gamma_{mkl}}{\partial\,h_i^{\,\nu}} - \frac{\partial}{\partial\,x^\varrho}\,\frac{\partial\,\gamma_{mkl}}{\partial\,h^{\,\nu}_{i,\varrho}} \right\}, \\ &\text{mit } W = \frac{1}{h'}\,\mathfrak{W}. \end{aligned} \right.$$

Beschränken wir uns auf die spezielle Relativität, indem wir

$$h_i^{\,\nu} = e_i\,h_{i,\nu} = \delta_{i\nu}$$

setzen. Dann wird (128c)[1])

1) Vgl. auch H. Tetrode, Ztschr. f. Phys. **49**. S. 858. 1928 Formeln (13) und (16), sowie den Text auf S. 862.

$$(129\,\mathrm{c})\quad \left\{\begin{aligned} T'_{i,\nu} &= \boldsymbol{R}\,\omega\,\psi^*\,\alpha_i\,\frac{\partial\psi}{\partial x^\nu} - \delta_{i\nu}\,W - e\,\psi^*\,\alpha_i\,\psi\,\varphi_\nu \\ &\quad + \boldsymbol{R}\,\frac{\omega}{4}\,e_\varrho\,e_\nu\,\frac{\partial}{\partial x^\varrho}\,(\psi^*\,\alpha_i\,\alpha_\nu\,\alpha_\varrho\,\psi), \\ &\text{mit} \\ W &= \boldsymbol{R}\,\omega\,e_\varrho\,\psi^*\,\alpha_\varrho\,\frac{\partial\psi}{\partial x^\varrho} - m\,c^2\,\psi^*\,\sigma\,\psi - e_\varrho\,e\,\psi^*\,\alpha_\varrho\,\psi\,\varphi_\varrho. \end{aligned}\right.$$

Insbesondere ist dann

$$T'_{44} = \boldsymbol{R}\,\omega\,\psi^*\,\alpha_{\bar\varrho}\,\frac{\partial\psi}{\partial x^{\bar\varrho}} - e\,\psi^*\,\alpha_{\bar\varrho}\,\psi\,\varphi_{\bar\varrho} + m\,c^2\,\psi^*\,\sigma\,\psi,$$

d. h. für den Energieoperator

$$(130\,\mathrm{c})\quad H = \alpha_{\bar\varrho}\left(\frac{h}{2\pi i}\,\frac{\partial}{\partial x^{\bar\varrho}} - \frac{e}{c}\,\varphi_{\bar\varrho}\right) + m\,c\,\sigma.$$

Ferner ist

$$T'_{4\bar\nu} = \boldsymbol{R}\,\omega\,\psi^*\,\frac{\partial\psi}{\partial x^{\bar\nu}} - e\,\psi^*\,\psi\,\varphi_{\bar\nu} + \boldsymbol{R}\,\frac{\omega}{4}\,\frac{\partial}{\partial x^{\bar\varrho}}\,(\psi^*\,\alpha_{\bar\nu}\,\alpha_{\bar\varrho}\,\psi);$$

setzen wir

$$(131)\qquad \alpha_1\,\alpha_2 = \mu_3$$

und zyklisch, so wird z. B.

$$\begin{aligned} T'_{41} &= \boldsymbol{R}\,\omega\,\psi^*\,\frac{\partial\psi}{\partial x^1} - e\,\psi^*\,\psi\,\varphi_1 \\ &\quad + \frac{\omega}{4}\left\{\frac{\partial}{\partial x^2}\,(\psi^*\,\mu_3\,\psi) - \frac{\partial}{\partial x^3}\,(\psi^*\,\mu_2\,\psi)\right\}. \end{aligned}$$

Der Impulsoperator lautet danach:

$$(132\,\mathrm{c})\qquad p_{\bar\nu} = \frac{h}{2\pi i}\,\frac{\partial}{\partial x^{\bar\nu}} - \frac{e}{c}\,\varphi_{\bar\nu};$$

andererseits bekommt man für den Drehimpuls:

$$\begin{aligned} M_1 &= x^2\,T'_{43} - x^3\,T'_{42} \\ &= \boldsymbol{R}\,\omega\,\psi^*\left(x^2\,\frac{\partial}{\partial x^3} - x^3\,\frac{\partial}{\partial x^2}\right)\psi - c\,\psi^*\,\psi\,[x^2\,\varphi_3 - x^3\,\varphi_2) \\ &\quad + \frac{\omega}{4}\left\{x^2\,\frac{\partial}{\partial x^1}\,(\psi^*\,\mu_2\,\psi) - x^2\,\frac{\partial}{\partial x^2}\,(\psi^*\,\mu_1\,\psi) - x^3\,\frac{\partial}{\partial x^3}\,(\psi^*\,\mu_1\,\psi)\right. \\ &\quad \left. + x^3\,\frac{\partial}{\partial x^1}\,(\psi^*\,\mu_3\,\psi)\right\}, \end{aligned}$$

folglich für den entsprechenden Operator

$$(133c)\quad \begin{cases} M_1 = \frac{h}{2\pi i}\left(x^2 \frac{\partial}{\partial x^3} - x^3 \frac{\partial}{\partial x^2}\right) \\ \qquad\qquad - \frac{e}{c}(x^2 \varphi_3 - x^3 \varphi_2) + \frac{i\mu_1}{2}\,. \end{cases}$$

2. Im Vorangehenden haben wir für die Vierbeine $h_{i,\nu}$ die Bose-Einsteinsche Statistik angenommen, d. h. die Klammersymbole in den V.-R. mit dem Minuszeichen gewählt. Nun fragt es sich, ob es möglich wäre, auch die Fermische Statistik auf die Vierbeine anzuwenden. Das Kriterium für die Zulässigkeit der V.-R. mit dem Pluszeichen ist folgendes (vgl. H.P. I, S. 29): Es sollen die *gewöhnlichen* Klammersymbole (mit dem Minuszeichen) $[\mathfrak{G}_\mu, Q_a]$, $[\mathfrak{G}_\mu, \mathfrak{P}^a]$ denselben Wert behalten, wenn man in $[Q_a, Q_\beta]$, $[\mathfrak{P}^a, \mathfrak{P}^\beta]$ und $[Q^a, \mathfrak{P}^\beta]$ das Minuszeichen durch das Pluszeichen ersetzt.

Nach diesem Kriterium ist aber die Frage bezüglich der Vierbeine zu *verneinen*. Denn man sieht aus der Form der Hamiltonfunktion (quadratisch in den $\mathfrak{P}^a$), daß $[\bar{\mathfrak{H}}_0, Q_a]$ beim Übergang vom Plus- zum Minuszeichen eine Änderung erfährt: die zum in $\mathfrak{P}^a$ quadratischen Anteil von $\mathfrak{H}_0$ gehörigen Klammersymbole sind nämlich in beiden Fällen verschieden und die Differenzen kompensieren sich nicht.

3. Das reine (Vakuum-)Gravitationsfeld ließe sich durch die $g_{\mu\nu}$ statt durch die $h_{i,\nu}$ beschreiben. Dann hätten wir mit einer etwas anderen Abart des „zweiten Falles" zu tun und bekämen wegen der allgemeinen Kovarianzgruppe vier Identitäten von der Gestalt $(\mathfrak{P}^a + \mathfrak{R}^{a4})\, c^4_{a\,r} = 0$.

Zusammenfassung

1. Verhält sich die Lagrangefunktion $\mathfrak{L}(Q_a; \dot{Q}_a)$ gegenüber der Gruppe[1])

$$(2')\quad \begin{cases} \delta x^\nu = a_r^{\nu,0}(x)\, \xi^r(x), \\ \delta Q_a = c^0_{a\,r}(x, Q)\, \xi^r + c^\sigma_{a\,r} \frac{\partial \xi^r}{\partial x^\sigma} \end{cases}$$

1) Der Übersichtlichkeit halber spezialisieren wir hier die Formeln auf den physikalisch wichtigen Fall $j = 1$.

wie eine skalare Dichte, so bestehen zwischen den Q und konjugierten Impulsen $\mathfrak{P}$ die Identitäten

$$\mathfrak{F}_r \equiv \underline{\mathfrak{P}^a c^4_{a\,r}} = 0\,. \tag{29'}$$

Falls nicht $\mathfrak{L}$, sondern $\mathfrak{L} + \mathfrak{L}'$ eine skalare Dichte ist, wobei $\mathfrak{L}'$ die zweiten Ableitungen der Q_a linear enthält, so tritt überall $\mathfrak{P}^a + \mathfrak{N}^{a4}$ an Stelle von $\mathfrak{P}^a$.

2. Infolgedessen ergibt die Auflösung der Gleichungen

$$\mathfrak{P}^a = \frac{\partial\,\mathfrak{L}}{\partial\,\dot{Q}_a}$$

nach den $\dot{Q}_a$:

$$\dot{Q}_a = \dot{Q}^0_a(\mathfrak{P}, Q) + \lambda^r c^4_{a\,r}\,, \tag{31'}$$

mit willkürlichen Raumzeitfunktionen λ^r.

Die Hamiltonfunktion nimmt sodann die Form

$$\mathfrak{H} = \mathfrak{H}_0(\mathfrak{P}, Q) + \lambda^r\,\mathfrak{F}_r \tag{35'}$$

an. Die Grundgleichungen der Theorie sind die kanonischen Feldgleichungen, die kanonischen V.-R., die Nebenbedingungen

$$\mathfrak{F}_r = 0 \quad \text{und} \quad \frac{d\,\mathfrak{F}_r}{d\,x^4} = 0\,.$$

3. Die infinitesimale Transformation der Gruppe läßt sich darstellen durch

$$\omega\,\delta^*\Phi = [\overline{\mathfrak{M}}, \Phi], \tag{45}$$

$$\mathfrak{M} = \underline{\mathfrak{P}^a\,\delta\,Q_a} - \mathfrak{G}_\mu\,\delta\,x^\mu\,. \tag{46}$$

[Φ beliebiges, nur von Q und $\mathfrak{P}$ abhängiges Funktional; $\mathfrak{G}_\mu$ Impulsenergie(pseudo)dichte].

Ein Spezialfall von $\overline{\mathfrak{M}}$ auf einem beliebigen Schnitt $x^4 = x_0{}^4$ ist $\overline{\varepsilon^r\,\mathfrak{F}_r}$. Daraus folgt, daß die $\mathfrak{F}_r$ auf Grund von $\mathfrak{F}_r = 0$ untereinander vertauschbar, d. h. die Nebenbedingungen $\mathfrak{F}_r = 0$ verträglich sind.

Ferner ist vermöge der Feldgleichungen

$$\frac{d\,\overline{\mathfrak{M}}}{d\,x^4} = 0\,, \tag{58}$$

woraus folgt

$$\overline{\mathfrak{M}} = \int d\,x^1\,d\,x^2\,d\,x^3 \left\{ \mathfrak{F}_r\,\frac{\partial\,\xi^r}{\partial\,x^4} - \frac{d\,\mathfrak{F}_r}{d\,x^4}\,\xi^r \right\} \tag{63'}$$

und

$$\frac{d^2 \mathfrak{F}_r}{(d x^4)^2} = 0 \tag{64'}$$

auf Grund der Feldgleichungen (zeitliche Fortpflanzung der Nebenbedingungen).

4. Das Schema der Grundgleichungen bleibt gegenüber der Gruppe invariant.

5. Als Beispiele werden das elektromagnetische Feld, das Diracsche Materiefeld und das Gravitationsfeld samt deren Wechselwirkungen behandelt. Die dabei in Betracht kommenden Gruppen sind die Eichinvarianzgruppe, die echte Beinkovarianzgruppe und die Gruppe der allgemeinen Relativitätstheorie.

Was insbesondere die Gravitation betrifft, so ist es unmöglich, die betreffenden Feldgrößen gemäß der Fermischen Statistik zu quanteln.

Hrn. Prof. Pauli spreche ich meinen aufrichtigen Dank aus für die Anregung zu dieser Arbeit und seine wertvollen Ratschläge.

Zürich, Physik. Institut der Eidgen. Technischen Hochschule, den 5. März 1930.

(Eingegangen 18. März 1930)

Chapter 19
Léon Rosenfeld (1930): Über die Gravitationswirkungen des Lichts

Léon Rosenfeld (1930). Über die Gravitationswirkungen des Lichts. *Zeitschrift für Physik*, 65: 589–599.

Über die Gravitationswirkungen des Lichtes.

Von **L. Rosenfeld,** zurzeit in Zürich.

(Eingegangen am 26. September 1930.)

Es wird das von einem elektromagnetischen Felde erzeugte Gravitationsfeld quantenmechanisch berechnet und gezeigt, daß die so entstandene Gravitationsenergie unendlich groß herauskommt, was eine neue Schwierigkeit für die Heisenberg-Paulische Quantelungstheorie der Wellenfelder bedeutet. Ferner werden die in erster Näherung möglichen Übergangsprozesse, an denen sich Licht- und Gravitationsquanten beteiligen, kurz erörtert.

Das Auftreten einer unendlich großen Selbstenergie des Elektrons bereitet bekanntlich* der Quantenelektrodynamik ernste Schwierigkeiten. Heisenberg hat die Frage aufgeworfen, ob nicht etwa schon unabhängig von jedem materiellen Einfluß, bei den Gravitationswirkungen des Lichts, analoge Verhältnisse herrschen. Die Antwort läßt sich nicht ohne weiteres durch Vergleich mit dem Verhalten des Elektrons erraten, da hier die Retardierung nicht vernachlässigt werden darf. Vorliegende Arbeit befaßt sich mit der Untersuchung dieser Frage.

*1. Das von einem elektromagnetischen Felde erzeugte Gravitationsfeld in erster Näherung***. Bezeichnet $\varkappa = 8\pi f/c^4$ die Einsteinsche Gravitationskonstante (f = Newtonsche Konstante), so wollen wir annehmen, daß die in Betracht kommenden Gravitationsfelder so wenig von dem Minkowskischen abweichen, daß wir sie nach Potenzen von $\varepsilon = \sqrt{\varkappa}$ entwickeln können und nur die in ε linearen Glieder zu berücksichtigen brauchen. In kartesischen Koordinaten x^1, x^2, x^3, $x^4 = ict$ können wir also schreiben:

$$g_{ik} = \delta_{ik} + \varepsilon\gamma_{ik}. \tag{1}$$

Setzen wir

$$\gamma = \sum_i \gamma_{ii} \tag{2}$$

und

$$\gamma'_{ik} = \gamma_{ik} - \tfrac{1}{2}\delta_{ik}\gamma, \tag{3}$$

* Vgl. W. Heisenberg u. W. Pauli, ZS. f. Phys. **56**, 1, 1929; R. Oppenheimer, Phys. Rev. **35**, 461, 1930; I. Waller, ZS. f. Phys. **62**, 673, 1930.

** Vgl. A. Einstein, Berl. Ber. 1918, S. 154, oder auch W. Pauli, Relativitätstheorie, Berlin 1921, Nr. 60, S. 736.

woraus umgekehrt

$$\left.\begin{aligned} \gamma' &= \sum_i \gamma'_{ii} = -\gamma, \\ \gamma_{ik} &= \gamma'_{ik} - \tfrac{1}{2}\delta_{ik}\gamma' \end{aligned}\right\} \qquad \begin{matrix}(4)\\(5)\end{matrix}$$

folgt, so legen wir das Koordinatensystem fest durch die Forderung*:

$$\frac{\partial \gamma'_{ik}}{\partial x^k} = 0. \tag{6}$$

Das vom elektromagnetischen Feld F_{ik} mit dem Maxwellschen Spannungstensor

$$\left.\begin{aligned} S_{ik} &= F_{ir}F_{kr} - \tfrac{1}{4}\delta_{ik}F, \\ F &= F_{rs}F_{rs}, \end{aligned}\right\} \tag{7}$$

erzeugte Gravitationsfeld wird gegeben durch die Gleichung

$$\sum_r \frac{\partial^2 \gamma'_{ik}}{(\partial x^r)^2} = -2\varepsilon S_{ik}. \tag{8}$$

Infolge von $\sum\limits_i S_{ii} = 0$ ist übrigens

$$\sum_i \frac{\partial^2 \gamma'}{(\partial x^i)^2} = 0, \tag{9}$$

so daß wir statt (8) auch

$$\sum_r \frac{\partial^2 \gamma_{ik}}{(\partial x^r)^2} = -2\varepsilon S_{ik} \tag{10}$$

schreiben können.

Dazu kommen die durch die Gravitationsglieder etwas modifizierten Maxwellschen Gleichungen, die wir nicht explizit hinzuschreiben brauchen.

Die entsprechende Lagrangefunktion lautet:

$$L = -\frac{1}{4}F - \frac{1}{8}\left(\frac{\partial \gamma'_{rs}}{\partial x^i}\frac{\partial \gamma'_{rs}}{\partial x^i} - \frac{1}{2}\frac{\partial \gamma'}{\partial x^i}\frac{\partial \gamma'}{\partial x^i}\right) + \frac{\varepsilon}{2}\gamma'_{rs}S_{rs}. \tag{11}$$

Daraus folgt für die Hamiltonfunktion, die wir in den Variablen q, q' (und nicht q, p) schreiben:

$$\mathfrak{H} = \mathfrak{H}_L + \mathfrak{H}_G + \mathfrak{W}, \tag{12}$$

wobei $\mathfrak{H}_L$ die gewöhnliche elektromagnetische Energiedichte $\frac{1}{2}(\mathfrak{E}^2 + \mathfrak{H}^2)$, $\mathfrak{H}_G$ den reinen Gravitationsanteil

$$\mathfrak{H}_G = \frac{1}{8}\left(\frac{\partial \gamma'_{rs}}{\partial x^i}\frac{\partial \gamma'_{rs}}{\partial x^i} - \frac{1}{2}\frac{\partial \gamma'}{\partial x^i}\frac{\partial \gamma'}{\partial x^i}\right) - \frac{1}{4}\left(\frac{\partial \gamma'_{rs}}{\partial x^4}\frac{\partial \gamma'_{rs}}{\partial x^4} - \frac{1}{2}\frac{\partial \gamma'}{\partial x^4}\cdot\frac{\partial \gamma'}{\partial x^4}\right), \tag{13}$$

* Die übliche Regel vom Weglassen des Summenzeichens wird überall befolgt, wo sie die Deutlichkeit nicht beeinträchtigt.

und $\mathfrak{W}$ den Wechselwirkungsanteil

$$\mathfrak{W} = \frac{\varepsilon}{2}\left(-\gamma'_{rs} S_{rs} + 2\gamma'_{4s} S_{4s} + 2\gamma'_{is} F_{4i} F_{4s} + \tfrac{1}{2}\gamma'_{44} F - \gamma' F_{4i} F_{4i}\right) \tag{14}$$

darstellt *.

Bei der Quantelung der γ'_{ik} wollen wir die Nebenbedingung (6) nach der „Fermischen Methode“* berücksichtigen. D. h. wir geben uns diese Bedingung sowie ihre zeitliche Ableitung auf einem Schnitt $t = \text{const}$ vor; dabei ist zu beachten, daß sie nicht als q-Zahlrelationen aufgefaßt werden dürfen. Es ist dann leicht zu sehen, mit Rücksicht auf den Erhaltungssatz $\partial S_{ik}/\partial x^k = 0$, daß sich die betrachteten Nebenbedingungen auf Grund der Feldgleichungen (8) mit der Zeit fortpflanzen.

Das elektromagnetische Feld (nullter Näherung) behandeln wir nach der zweiten Heisenberg-Paulischen Methode*; da wir keine Materie in Betracht ziehen, haben wir es nur mit den transversalen Eigenschwingungen zu tun; wir zerlegen die Feldstärken nach fortschreitenden Wellen; als Randbedingung wählen wir eine zyklische Bedingung mit der Periode L und lassen zum Schluß L gegen unendlich streben **. Wir setzen also

$$\left.\begin{aligned}
\mathfrak{E} &= \frac{\alpha}{\sqrt{L^3}} \sum_{\mathfrak{k},\lambda} \sqrt{\frac{k}{L}}\, \mathfrak{e}^{\mathfrak{k}\lambda} \left(A_{\mathfrak{k}\lambda}\, e^{\frac{2\pi i \mathfrak{k}\mathfrak{r}}{L}} - A_{-\mathfrak{k},-\lambda}\, e^{-\frac{2\pi i \mathfrak{k}\mathfrak{r}}{L}}\right),\\
\mathfrak{H} &= \frac{\alpha}{\sqrt{L^3}} \sum_{\mathfrak{k},\lambda} \sqrt{\frac{k}{L}}\, \mathfrak{h}^{\mathfrak{k}\lambda} \left(A_{\mathfrak{k}\lambda}\, e^{\frac{2\pi i \mathfrak{k}\mathfrak{r}}{L}} - A_{-\mathfrak{k},-\lambda}\, e^{-\frac{2\pi i \mathfrak{k}\mathfrak{r}}{L}}\right);
\end{aligned}\right\} \tag{15}$$

dabei ist der Normierungsfaktor $\alpha = \frac{1}{i}\sqrt{\frac{ch}{2}}$; der Index $\lambda = 1{,}2$ bezeichnet die beiden zueinander senkrechten Eigenschwingungen mit dem Ausbreitungsvektor $\mathfrak{k}$, vom Betrage $k = |\mathfrak{k}|$; $\mathfrak{e}^{\mathfrak{k}\lambda}$ und $\mathfrak{h}^{\mathfrak{k}\lambda}$ sind zueinander und zu $\mathfrak{k}$ senkrechte Einheitsvektoren, und zwar ist $\mathfrak{h}^{\mathfrak{k}\lambda} = \mathfrak{e}^{\mathfrak{k}\lambda} \times \mathfrak{k}/k$, ferner $\mathfrak{e}^{-\mathfrak{k},\lambda} = \mathfrak{e}^{\mathfrak{k},\lambda}$, $\mathfrak{h}^{-\mathfrak{k},\lambda} = -\mathfrak{h}^{\mathfrak{k},\lambda}$; endlich hängen die Amplituden A mit den Anzahl- und Phasenvariablen folgendermaßen zusammen:

$$A_{\mathfrak{k}\lambda} = e^{-\frac{2\pi i}{h}\Theta_{\mathfrak{k}\lambda}}\, N_{\mathfrak{k}\lambda}^{1/2}, \quad A_{-\mathfrak{k},-\lambda} = N_{\mathfrak{k}\lambda}^{1/2}\, e^{\frac{2\pi i}{h}\Theta_{\mathfrak{k}\lambda}}. \tag{16}$$

Der Übersichtlichkeit halber werden wir oft im folgenden einen Zustand $(\mathfrak{k}_r, \lambda_r)$ durch den Buchstaben r bezeichnen und dementsprechend

* Dabei ist Φ_4 gleich Null gewählt; vgl. W. Heisenberg u. W. Pauli, ZS. f. Phys. **59**, 171 ff., 1930.

** L. Landau u. R. Peierls, ZS. f. Phys. **62**, 197, 1930.

$A_{\mathfrak{k}_r \lambda_r} \equiv A_r$, $A_{-\mathfrak{k}_r, -\lambda_r} \equiv B_r$ setzen. Ferner führen wir noch folgende Abkürzungen ein:

$$\left.\begin{aligned}
I_{\pm 5}(rs) &= \frac{1}{L^3} e^{\frac{2\pi i}{L}(\mathfrak{k}_r \pm \mathfrak{k}_s)\mathfrak{r}}, \\
I_{\pm 6}(rs) &= \frac{1}{L^3} \frac{e^{\frac{2\pi i}{L}(\mathfrak{k}_r \pm \mathfrak{k}_s)\mathfrak{r}}}{|\mathfrak{k}_r \pm \mathfrak{k}_s|^2 - (k_r \pm k_s)^2} \cdot \frac{L^2}{2\pi^2} \\
&= \pm \frac{L^2}{4\pi^2} \frac{1}{k_r k_s (\cos\Theta_{rs} - 1)} I_{\pm 5}(rs) \\
&\quad \left(\cos\Theta_{rs} = \frac{\mathfrak{k}_r \mathfrak{k}_s}{k_r k_s}\right), \\
I_{\pm i}(rs) &= \frac{\partial I_{\pm 6}(rs)}{\partial x^i} = \frac{2\pi i}{L} (\mathfrak{k}_r \pm \mathfrak{k}_s)_i I_{\pm 6}(rs) \quad (i = 1, 2, 3), \\
I_{\pm 4}(rs) &= -\frac{2\pi}{L} (k_r \pm k_s) I_{\pm 6}(rs), \\
I^*_{\pm 4}(rs) &= +\frac{2\pi}{L} (k_r \pm k_s) I^*_{\pm 6}(rs),
\end{aligned}\right\} \tag{17}$$

(x^* = komplex konjugiertes von x).

Diese Erklärungen gelten nur, so lange $\mathfrak{k}_r$ nicht parallel $\mathfrak{k}_s$ ist; wie wir sehen werden, scheidet dieser singuläre Fall von selbst im Laufe der folgenden Rechnungen aus. Schließlich definieren wir noch den Tensor $\mathfrak{s}^{rs}_{ik}$ durch folgende Gleichungen:

$$\left.\begin{aligned}
-\mathfrak{s}^{rs}_{44} &\equiv \mathfrak{w} = \tfrac{1}{2} \sum (\mathfrak{e}^r_i \mathfrak{e}^s_i + \mathfrak{h}^r_i \mathfrak{h}^s_i), \\
\mathfrak{s}^{rs}_{il} &= \delta_{il} \mathfrak{w} - \tfrac{1}{2} (\mathfrak{e}^r_i \mathfrak{e}^s_l + \mathfrak{h}^r_i \mathfrak{h}^s_l) - \tfrac{1}{2} (\mathfrak{e}^s_i \mathfrak{e}^r_l + \mathfrak{h}^s_i \mathfrak{h}^r_l), \\
\mathfrak{s}^{rs}_{i4} &= \mathfrak{s}^{rs}_{4i} = \frac{i}{2} [(\mathfrak{e}^r \times \mathfrak{h}^s)_l + (\mathfrak{e}^s \times \mathfrak{h}^r)_l] \quad (i, l = 1, 2, 3).
\end{aligned}\right\} \tag{18}$$

Nach (15) und (7) haben wir mit diesen Bezeichnungen

$$S_{ik} = \alpha^2 \sum_{rs} \sqrt{\frac{k_r k_s}{L^2}} \mathfrak{s}^{rs}_{ik} \{A_r A_s I_{+5}(rs) + B_r B_s I^*_{+5}(rs) - A_r B_s I_{-5}(rs) - B_r A_s I^*_{-5}(rs)\}; \tag{19}$$

wenn wir annehmen, daß nur das durch das Lichtfeld (19) erzeugte Gravitationsfeld vorhanden ist, so wird es mit Rücksicht auf die Phasen der A, B und auf die zyklische Randbedingung gegeben durch folgende Lösung von (10):

$$\gamma_{ik} = \varepsilon\alpha^2 \sum_{rs} \sqrt{\frac{k_r k_s}{L^2}}\, \mathfrak{s}_{ik}^{rs} \{A_r A_s I_{+6}(rs) + B_r B_s I_{+6}^*(rs) - A_r B_s I_{-6}(rs) - B_r A_s I_{-6}^*(rs)\}; \tag{20}$$

diese auch für q-Zahlen gültige Lösung erfüllt die Nebenbedingung (6) wegen des Erhaltungssatzes

$$\frac{\partial S_{ik}}{\partial x^k} = 0;$$

ferner sieht man leicht ein, daß

$$\lim_{L \to \infty} I_{-6}(rr) = 0 \tag{21}$$

ist, wenn man bedenkt, daß es das (retardierte) Potential einer Belegung von konstanter, in Limes überall verschwindender Dichte $1/L^3$ darstellt.

2. *Berechnung der Gravitationsenergie.* Da das Gravitationsfeld (20) von der ersten Ordnung in ε ist, so ist die Gravitationsenergie $\int(\mathfrak{H}_G + \mathfrak{W})\,dV$ nach (13) und (14) von der Ordnung ε^2. Die im Sinne der Störungsrechnung richtige Störung (zweiter Ordnung) der Energie $\overline{\mathfrak{H}}_L = \sum_r (N_r + \frac{1}{2}) h\nu_r$ eines durch die Zahlen N_r der Lichtquanten von der Art r $\left(\nu_r = \frac{k_r c}{L}\right)$ charakterisierten Zustandes des Lichtfeldes bekommt man also, indem man in $\mathfrak{H}_G + \mathfrak{W}$ für die γ_{ik} die Ausdrücke (20) einsetzt und das zum betreffenden Zustand zugehörige Diagonalglied berechnet.

Eine Vereinfachung tritt dabei zunächst dadurch ein, daß für das Feld (20) $\gamma = -\gamma' = 0$ und mithin $\gamma'_{ik} = \gamma_{ik}$ ist. Die übrigbleibenden Glieder haben die Form

$$\begin{aligned}\varepsilon^2\alpha^4 \sum_{rsmn} \sqrt{\frac{k_r k_s k_m k_n}{L^4}}\, \mathfrak{s}_{ik}^{rs}\mathfrak{s}_{ik}^{mn} \int \{&A_r A_s I_{+\tau}(rs) + B_r B_s I_{+\tau}^*(rs) \\ &- A_r B_s I_{-\tau}(rs) - B_r A_s I_{-\tau}^*(rs)\} \\ \cdot \{A_m A_n I_{+\varrho}(mn) + B_m B_n I_{+\varrho}^*(mn) &- A_m B_n I_{-\varrho}(mn) - B_m A_n I_{-\varrho}^*(mn)\}\,dV \\ &(\varrho, \tau = 1, 2, \cdots, 6)\end{aligned} \tag{22}$$

oder eine ähnliche, wo $\mathfrak{s}_{ik}^{rs}\mathfrak{s}_{ik}^{mn}$ zu ersetzen ist durch

$$\mathfrak{s}_{i4}^{rs}\mathfrak{s}_{i4}^{mn} - \tfrac{1}{2}\mathfrak{s}_{ik}^{rs}(\mathfrak{e}_i^m \mathfrak{e}_k^n + \mathfrak{e}_i^n \mathfrak{e}_k^m), \tag{23}$$

bzw.

$$(\mathfrak{h}^r \mathfrak{h}^s + \mathfrak{e}^r \mathfrak{e}^s)(\mathfrak{h}^m \mathfrak{h}^n - \mathfrak{e}^m \mathfrak{e}^n). \tag{23'}$$

Im Integrand von (22) enthalten nur Produkte mit zwei Faktoren A und zwei Faktoren B mit denselben Indizes Diagonalglieder. Wenn man noch (21) berücksichtigt, reduziert sich (22) auf

$$\varepsilon^2 \alpha^4 \sum_{rs} \frac{k_r k_s}{L^2} \sum_{ik} (\mathfrak{s}_{ik}^{rs})^2 \int dV \{2 A_r A_s B_r B_s I_{+\tau}(rs) I^*_{+\varrho}(rs)$$
$$+ 2 B_r B_s A_r A_s I^*_{+\tau}(rs) I_{+\varrho}(rs) + A_r B_s A_s B_r I_{-\tau}(rs) I_{-\varrho}(sr)$$
$$+ A_r B_s B_r A_s I_{-\tau}(rs) I^*_{-\varrho}(rs) + B_r A_s A_r B_s I^*_{-\tau}(rs) I_{-\varrho}(rs)$$
$$+ B_r A_s B_s A_r I^*_{-\tau}(rs) I^*_{-\varrho}(sr)\}. \tag{24}$$

Um nun die Ausdrücke $\sum_{i,k} (\mathfrak{s}_{ik}^{rs})^2$ und (23) bequem auszuwerten, bemerken wir, daß sie kovariant gegen Drehungen sind. Wählen wir also $\mathfrak{h}^s$, $\mathfrak{e}^s$, $\mathfrak{k}_s$ als Achsenkreuz $0xyz$, so reduziert sich der Tensor $\mathfrak{s}_{ik}^{rs}$ nach (18) zu

$$\left|\begin{array}{cccc}
\frac{1}{2}(\mathfrak{e}_2 - \mathfrak{h}_1), & -\frac{1}{2}(\mathfrak{e}_1 + \mathfrak{h}_2), & -\frac{1}{2}\mathfrak{h}_3, & \frac{i}{2}\mathfrak{h}_3, \\
-\frac{1}{2}(\mathfrak{e}_1 + \mathfrak{h}_2), & -\frac{1}{2}(\mathfrak{e}_2 - \mathfrak{h}_1), & -\frac{1}{2}\mathfrak{e}_3, & \frac{i}{2}\mathfrak{e}_3, \\
-\frac{1}{2}\mathfrak{h}_3, & -\frac{1}{2}\mathfrak{e}_3, & \frac{1}{2}(\mathfrak{e}_2 + \mathfrak{h}_1), & -\frac{i}{2}(\mathfrak{e}_2 + \mathfrak{h}_1), \\
\frac{i}{2}\mathfrak{h}_3, & \frac{i}{2}\mathfrak{e}_3, & -\frac{i}{2}(\mathfrak{e}_2 + \mathfrak{h}_1), & -\frac{1}{2}(\mathfrak{e}_2 + \mathfrak{h}_1),
\end{array}\right|$$

woraus folgt, unter Benutzung der Orthogonalitätsrelationen, für

$$\sum_{i,k} (\mathfrak{s}_{ik}^{rs})^2 \colon \ \frac{1}{2}\left(1 - \frac{\mathfrak{k}_3}{k}\right)^2,$$
$$\sum_i (\mathfrak{s}_{i4}^{rs})^2 - \frac{1}{2}\mathfrak{s}_{ik}^{rs}(\mathfrak{e}_i^r \mathfrak{e}_k^s + \mathfrak{e}_i^s \mathfrak{e}_k^r) \colon \ \frac{1}{4}\left(1 - \frac{\mathfrak{k}_3}{k}\right)^2,$$

d. h. wenn man zum ursprünglichen Koordinatensystem zurückkehrt,

$$\left.\begin{array}{l}
\sum_{i,k} (\mathfrak{s}_{ik}^{rs})^2 = \frac{1}{2}(1 - \cos\Theta_{rs})^2, \\
\sum_i (\mathfrak{s}_{i4}^{rs})^2 - \frac{1}{2}\mathfrak{s}_{ik}^{rs}(\mathfrak{e}_i^r \mathfrak{e}_k^s + \mathfrak{e}_i^s \mathfrak{e}_k^r) = \frac{1}{4}(1 - \cos\Theta_{rs})^2.
\end{array}\right\} \tag{25}$$

Was den Ausdruck (23′) betrifft, wollen wir zunächst festsetzen, daß bei festem Achsenkreuz $\mathfrak{h}^s$, $\mathfrak{e}^s$, $\mathfrak{k}_s$ und festem λ_s, die (bisher willkürliche) Richtung von $\mathfrak{e}^r$ bzw. $\mathfrak{h}^r$ jeweils mit der Schnittlinie der Ebenen ($\mathfrak{e}^s$, $\mathfrak{h}^s$) und ($\mathfrak{e}^r$, $\mathfrak{h}^r$) zusammenfällt, je nachdem $\lambda_r = \lambda_s$ ist oder nicht; dementsprechend setzen wir

$$\mathfrak{e}^r \mathfrak{e}^s = \cos\varphi_{rs}, \quad \text{wenn } \lambda_r = \lambda_s;$$

dann wird bei gegebenem $\mathfrak{k}_r$

$$\left.\begin{aligned} (\mathfrak{h}^r \mathfrak{h}^s)^2 - (\mathfrak{e}^r \mathfrak{e}^s)^2 &= -\cos^2 \varphi_{rs} \sin^2 \Theta_{rs}, \quad \text{wenn } \lambda_r = \lambda_s, \\ (\mathfrak{h}^r \mathfrak{h}^s)^2 - (\mathfrak{e}^r \mathfrak{e}^s)^2 &= +\sin^2 \varphi_{rs} \sin^2 \Theta_{rs}, \quad \text{wenn } \lambda_r \neq \lambda_s. \end{aligned}\right\} \tag{25'}$$

Jetzt sieht man, wie die Singularitäten in den $I_{\pm\varrho}(rs)$ für $\Theta_{rs} = 0$ durch die Faktoren (25), (25′) aufgehoben werden.

Für den von $\mathfrak{H}_G$ herrührenden Anteil der Störungsenergie bekommt man nach (13), (24), (25)

$$\begin{aligned} \overline{\mathfrak{H}}_G = \frac{\varepsilon^2 \alpha^4}{32\pi^2} \cdot \frac{1}{L^3} \sum_{rs} \Big[(\cos\Theta_{rs} - 1) \{ 2 A_r A_s B_r B_s + 2 B_r B_s A_r A_s \\ - (A_r B_s A_s B_r + A_r B_s B_r A_s + B_r A_s A_r B_s + B_r A_s B_s A_r) \} \\ + \frac{(k_r + k_s)^2}{k_r k_s} \{ 2 A_r A_s B_r B_s + 2 B_r B_s A_r A_s \} \\ + \frac{(k_r - k_s)^2}{k_r k_s} (A_r B_s A_s B_r + A_r B_s B_r A_s + B_r A_s A_r B_s + B_r A_s B_s A_r) \Big], \end{aligned}$$

oder, nach gesonderter Behandlung der Glieder mit $r = s$,

$$\begin{aligned} \overline{\mathfrak{H}}_G = \frac{\varepsilon^2 \alpha^4}{16\pi^2} \cdot \frac{1}{L^3} \sum_{rs} \Big[(\cos\Theta_{rs} + 1)(A_r B_r - B_r A_r)(A_s B_s - B_s A_s) \\ + \frac{k_r^2 + k_s^2}{k_r k_s} (A_r B_r + B_r A_r)(A_s B_s + B_s A_s) \Big] \\ - \frac{\varepsilon^2 \alpha^4}{8\pi^2} \cdot \frac{1}{L^3} \sum_r \left[(A_r B_r + B_r A_r)^2 + (A_r B_r - B_r A_r)^2 - 2 A_r^2 B_r^2 - 2 B_r^2 A_r^2 \right], \end{aligned}$$

oder schließlich gemäß (16)

$$\begin{aligned} \overline{\mathfrak{H}}_G = \frac{\varepsilon^2 \alpha^4}{16\pi^2} \cdot \frac{1}{L^3} \sum_{rs} (\cos\Theta_{rs} + 1) + \frac{\varepsilon^2 \alpha^4}{4\pi^2} \sum_r \frac{1}{L^3} \\ + \frac{\varepsilon^2 \alpha^4}{16\pi^2} \cdot \frac{1}{L^3} \sum_{rs} \frac{k_r^2 + k_s^2}{k_r k_s} (2 N_r + 1)(2 N_s + 1). \end{aligned} \tag{26}$$

Durch eine analoge, von (14), (24), (25) und (25′) ausgehende Rechnung findet man, wenn man noch berücksichtigt, daß die Mittelwerte von $\cos^2 \varphi_{rs}$ und $\sin^2 \varphi_{rs}$ gleich sind, daß der Wechselwirkungsanteil $\overline{\mathfrak{W}}$ gleich Null ist. Mithin bleibt für die gesuchte Störungsenergie der Ausdruck (26).

Hätten wir es mit einem *klassischen* Wellenpaket zu tun, so wären in (26) die $(2N + 1)$ durch $2N$ zu ersetzen und die erste Zeile zu streichen, und wir bekämen für die Gravitationsenergie einen *endlichen* Wert. Quantenmechanisch hingegen finden wir ein wegen des Auftretens von Schwingungen mit beliebig kurzer Wellenlänge unendlich großes Zusatzglied, welches übrigens noch dann bestehen bleibt, wenn man in (19) $A_r B_s$ durch $B_s A_r$ ersetzt, um die Nullpunktsenergie der Strahlung zu beseitigen. Dieses unendliche Glied setzt sich aus zwei Teilen zusammen: einem von den Lichtquantenzahlen unabhängigen und einem *ebenfalls unendlichen,* zu den Lichtquantenzahlen proportionalen Anteil.

Man kann die unendlichen Faktoren von der Gestalt

$$\lim_{L \to \infty} \frac{1}{L^3} \int k^n \, d\mathfrak{k}_1 \, d\mathfrak{k}_2 \, d\mathfrak{k}_3 \quad (n = 1, 2, 3)$$

auf eine andere, lehrreiche Form bringen, die zugleich zeigt, daß der Übergang zum Limes $L = \infty$ unwesentlich ist. Den Normierungsfaktor $1/L^3$ kann man nämlich schreiben, wenn $u(\mathfrak{k}, \mathfrak{r})$ eine normierte Eigenfunktion darstellt,

$$\frac{1}{L^3} = u(\mathfrak{k}, \mathfrak{r}) \, u^*(\mathfrak{k}, \mathfrak{r}) = \int u(\mathfrak{k}, \mathfrak{r}) \, u^*(\mathfrak{k}, \mathfrak{r}') \, \delta(\mathfrak{r} - \mathfrak{r}') \, dV';$$

dann ist

$$\frac{1}{L^3} \int k^n \, d\mathfrak{k}_1 \, d\mathfrak{k}_2 \, d\mathfrak{k}_3 = \int \delta(\mathfrak{r} - \mathfrak{r}') \, dV' \int k^n \, u(\mathfrak{k}, \mathfrak{r}) \, u^*(\mathfrak{k}, \mathfrak{r}') \, d\mathfrak{k}_1 \, d\mathfrak{k}_2 \, d\mathfrak{k}_3;$$

wegen

$$\delta(\mathfrak{r} - \mathfrak{r}') = \int u(\mathfrak{k}, \mathfrak{r}) \, u^*(\mathfrak{k}, \mathfrak{r}') \, d\mathfrak{k}_1 \, d\mathfrak{k}_2 \, d\mathfrak{k}_3$$

ist, mit einer Bezeichnung von Landau und Peierls (l. c. S. 189),

$$\begin{aligned} \frac{1}{L^3} \int k^n \, d\mathfrak{k}_1 \, d\mathfrak{k}_2 \, d\mathfrak{k}_3 &= \int \delta(\mathfrak{r} - \mathfrak{r}') \, dV' \, (-\Delta_{\mathfrak{r}})^{n/2} \, \delta(\mathfrak{r} - \mathfrak{r}') \\ &= [(-\Delta_{\mathfrak{r}})^{n/2} \, \delta(\mathfrak{r} - \mathfrak{r}')]_{\mathfrak{r} = \mathfrak{r}'}; \end{aligned}$$

man kann also auch sagen, daß die Unendlichkeit darauf beruht, daß man einem Lichtquant keinen endlichen Radius zuschreiben kann. Die Analogie mit dem Fall des Elektrons braucht kaum betont zu werden.

3. Übergangsprozesse erster Näherung. Um nun die durch die Wechselwirkung $\mathfrak{W}$ hervorgerufenen Übergänge zu übersehen, wollen wir zunächst neben den Lichtwellen (15) auch reine Gravitationswellen einführen, die wir zur nullten Näherung unseres Wellenfeldes mitrechnen.

Nach Einstein gibt es zweierlei Gravitationswellen im Vakuum. Mit Rücksicht auf die Nebenbedingung (6) lassen sie sich durch die Komponenten $\gamma_{11} - \gamma_{22}$ bzw. γ_{12} beschreiben, wenn der Ausbreitungsvektor als z-Achse gewählt ist. Wir können dann noch $\gamma_{11} = -\gamma_{22}$ setzen, um die Vereinfachung $\gamma = 0$ zu erreichen. Für ein beliebiges Paket solcher Wellen haben wir dann

$$\gamma = 0 \text{ und } \gamma_{i4} = 0; \tag{27}$$

die übrigen $\gamma_{\mu\nu}$ ($\mu, \nu = 1, 2, 3$) lauten, wenn $\{D^{(r)}_{\alpha\beta}\}$ diejenige Drehung darstellt, welche den Ausbreitungsvektor $\mathfrak{k}_r$ in die z-Richtung überführt,

$$\begin{aligned}\gamma_{\mu\nu} = \frac{1}{\pi}\sqrt{\frac{hc}{L^3}}\sum_{\mathfrak{k}_r}\sqrt{\frac{L}{k_r}}\Big\{\frac{1}{2}(D^{(r)}_{\mu 1}D^{(r)}_{\nu 1} - D^{(r)}_{\mu 2}D^{(r)}_{\nu 2})\Big(F_{\mathfrak{k}_r}\, e^{\frac{2\pi i\,\mathfrak{k}_r \mathfrak{r}}{L}} + F^{\dagger}_{-\mathfrak{k}_r} e^{-\frac{2\pi i\,\mathfrak{k}_r\mathfrak{r}}{L}}\Big) \\ + \tfrac{1}{2}(D^{(r)}_{\mu 1}D^{(r)}_{\nu 2} + D^{(r)}_{\mu 2}D^{(r)}_{\nu 1})\Big(G_{\mathfrak{k}_r}\, e^{\frac{2\pi i\,\mathfrak{k}_r\mathfrak{r}}{L}} + G^{\dagger}_{-\mathfrak{k}_r}\, e^{-\frac{2\pi i\,\mathfrak{k}_r\mathfrak{r}}{L}}\Big)\Big\};\end{aligned} \tag{28}$$

dabei ist, wenn $M_{\mathfrak{k}_r,1}$, $M_{\mathfrak{k}_r,2}$ die Anzahlen von Gravitationsquanten der ersten bzw. zweiten Art in der Richtung $\mathfrak{k}_r$ bezeichnen,

$$\left.\begin{aligned} F_{\mathfrak{k}_r} &= e^{-\frac{2\pi i}{h}\Theta_{\mathfrak{k}_r,1}} M^{1/2}_{\mathfrak{k}_r,1}, \qquad & F^{\dagger}_{-\mathfrak{k}_r} &= M^{1/2}_{\mathfrak{k}_r,1}\, e^{\frac{2\pi i}{h}\Theta_{\mathfrak{k}_r,1}}, \\ G_{\mathfrak{k}_r} &= e^{-\frac{2\pi i}{h}\Theta_{\mathfrak{k}_r,2}} M^{1/2}_{\mathfrak{k}_r,2}, \qquad & G^{\dagger}_{-\mathfrak{k}_r} &= M^{1/2}_{\mathfrak{k}_r,2}\, e^{\frac{2\pi i}{h}\Theta_{\mathfrak{k}_r,2}}. \end{aligned}\right\} \tag{29}$$

Die Energie dieser Gravitationswellen ist nach (13)

$$\overline{\mathfrak{H}}_G = \sum_{\mathfrak{k}_r}\{(M_{\mathfrak{k}_r,1} + \tfrac{1}{2}) + (M_{\mathfrak{k}_r,2} + \tfrac{1}{2})\}\, h\,\nu_r. \tag{30}$$

Gemäß (14), welches sich hier wegen (27) auf

$$\mathfrak{W} = \varepsilon\gamma_{\mu\nu}(F_{4\mu}F_{4\nu} - \tfrac{1}{2}S_{\mu\nu}) \tag{31}$$

reduziert, kommen in erster Näherung (d. h. mit einer ε^2 proportionalen Wahrscheinlichkeit) nur solche Übergänge vor, an denen ein Gravitationsquant und zwei Lichtquanten teilnehmen. Berücksichtigen wir (auf Grund wohlbekannter Überlegungen) nur diejenigen Prozesse, die unter Erhaltung der Gesamtenergie stattfinden, so muß gelten, da auch der Gesamtimpuls erhalten bleibt,

$$\left.\begin{aligned} k_r &= k_s + k_t, \\ \mathfrak{k}_r &= \mathfrak{k}_s + \mathfrak{k}_t, \end{aligned}\right|$$

woraus folgt, daß die drei am Prozeß beteiligten Quanten dieselbe Richtung haben müssen.

Bezeichnen wir mit t den Zustand des Gravitationsquants, mit r und s diejenigen der Lichtquanten, so sind nach (31) die Übergangswahrscheinlichkeiten pro Zeiteinheit von der Gestalt

$$\varepsilon^2 c^2 h\,\frac{k_r k_s}{k_t}\cdot\frac{1}{L^3}\, w_{rst} f(N_r, N_s, M_t); \tag{32}$$

dabei ist $f(N_r, N_s, M_t)$ das übliche Produkt von Faktoren N_r, N_s, M_t; $N_r + 1$, $N_s + 1$, $M_t + 1$; $N_r + 2$, $N_s + 2$, je nach dem Prozeß; wenn ferner Θ_{rt} der Winkel zwischen der Polarisationsrichtung $\mathfrak{e}^r$ des Lichtquants r und der ausgezeichneten y-Richtung in der Wellenebene des Gravitationsquants bedeutet, so ist

$$w_{rst} = \tfrac{1}{4}\cos^2 2\Theta_{rt},$$

wenn entweder das Gravitationsquant von der ersten Art und die beiden Lichtquanten gleich polarisiert ($\lambda_r = \lambda_s$) sind, oder das Gravitationsquant von der zweiten Art und die beiden Lichtquanten verschieden polarisiert ($\lambda_r \neq \lambda_s$) sind,

$$w_{rst} = \tfrac{1}{4}\sin^2 2\Theta_{rt},$$

in den beiden anderen möglichen Fällen.

Die Übergangsprozesse selber lassen sich folgendermaßen beschreiben:

1. Verschwinden eines Gravitationsquants und Entstehen zweier (verschiedener oder gleicher) Lichtquanten;

2. Verschwinden zweier Lichtquanten und Entstehen eines Gravitationsquants;

3. Verschwinden eines Lichtquants und Entstehen eines anderen Lichtquants und eines Gravitationsquants (Frequenzverkleinerung eines Lichtquants!);

4. Verschwinden eines Lichtquants und eines Gravitationsquants und Entstehen eines anderen Lichtquants (Frequenzvergrößerung eines Lichtquants!).

Denken wir uns also einen anfangs nur von Strahlung gefüllten Hohlraum (ohne Kohlestäubchen!), so genügen schon die Gravitationswirkungen erster Näherung zwischen den Lichtquanten, um das Plancksche Gleichgewicht (mit einer zu $1/\varkappa$ proportionalen Geschwindigkeit) herzustellen.

Herrn Prof. Pauli bin ich für viele kritische Bemerkungen und Ratschläge herzlich dankbar.

Zürich, Physikalisches Institut der Eidgenössischen Technischen Hochschule, 14. August 1930.

Nachtrag bei der Korrektur.

Statt wie im § 2 als Ausgangszustand Lichtquanten mit genau bekanntem Impuls zu nehmen, kann man auch die Berechnung des Mittelwertes von $\overline{\mathfrak{H}_G + \mathfrak{W}}$ für das allgemeinste Wellenpaket durchführen.

Die Anfangsverteilung des Zustandes r sei charakterisiert durch eine (komplexe) Eigenfunktion $\varphi_r(N_r)$ mit der Bedingung:

$$\sum_{N_r=0}^{\infty} |\varphi_r(N_r)|^2 = 1;$$

der Ausgangspunkt ist dann definiert durch die Angabe willkürlicher $\varphi_r(N_r)$ für alle r, mit der einzigen Einschränkung, daß die Gesamtzahl N der Lichtquanten eine gegebene Konstante ist:

$$\sum_r \sum_{N_r} N_r |\varphi_r(N_r)|^2 = N;$$

$$\varphi_r(N_r) = 0 \quad \text{für} \quad N_r > N.$$

Wir haben dann zu berechnen:

$$\text{Mittelwert von } \overline{(\mathfrak{H}_G + \mathfrak{W})}$$
$$= \sum_{N_0 N_1 \ldots} \varphi_0^*(N_0)\, \varphi_1^*(N_1) \ldots \overline{(\mathfrak{H}_G + \mathfrak{W})}\, \varphi_0(N_0)\, \varphi_1(N_1) \ldots$$

Für $N = 0$ (kein Lichtquant) bekommen wir natürlich dasselbe Resultat wie im § 2. Betrachten wir nun den Fall *eines* Lichtquants, d. h.

$$\varphi_r(N_r) = 0 \quad \text{für} \quad N_r > 1$$
$$|\varphi_r(N_r = 0)|^2 + |\varphi_r(N_r = 1)|^2 = 1,$$
$$\sum_r |\varphi_r(N_r = 1)|^2 = 1.$$

Zunächst kommen in Betracht die Diagonalglieder (26): der Beitrag dieser Glieder ist wiederum unendlich, wie im § 2. Ferner können, wie man leicht überlegt, die anderen Glieder nur endliche Beiträge liefern.

Chapter 20
Matvei Bronstein (1936): Quantentheorie schwacher Gravitationsfelder

Matvei Bronstein (1936). Quantentheorie Schwacher Gravitationsfelder. *Physikalische Zeitschrift der Sowjetunion*, 9: 140–157.

QUANTENTHEORIE SCHWACHER GRAVITATIONSFELDER.[1]

Von M. Bronstein.

(Eingegangen am 2. Januar 1936.)

§ 1. Allgemeines. § 2. Hamiltonsche Form und ebene Wellen. § 3. Vertauschungsrelationen und Eigenwerte der Energie. § 4. Ein wenig Gedankenexperimentieren! § 5. Wechselwirkung mit der Materie. § 6. Energieübertragung durch die Gravitationswellen. § 7. Herleitung des Newtonschen Gravitationsgesetzes.

§ 1. Allgemeines.

Die Abweichungen eines raumzeitlichen Kontinuums von der „Euklidizität" können bekanntlich durch die Komponenten des Riemann-Christoffelschen Tensors charakterisiert werden. Wenn diese Abweichungen klein sind, kann dieses Tensorfeld vierter Stufe aus einem symmetrischen Tensorfeld zweiter Stufe auf folgende Weise abgeleitet werden:

$$(\mu\rho\nu\sigma) = \frac{1}{2}\left(\frac{\partial^2 h_{\mu\nu}}{\partial x_\rho \partial x_\sigma} + \frac{\partial^2 h_{\rho\sigma}}{\partial x_\mu \partial x_\nu} - \frac{\partial^2 h_{\mu\sigma}}{\partial x_\rho \partial x_\nu} - \frac{\partial^2 h_{\rho\nu}}{\partial x_\mu \partial x_\sigma}\right), \tag{1}$$

wo $h_{\mu\nu}$ ist die kleine Abweichung des fundamentalen metrischen Tensors von seinem Minkowskischen Wert $\Delta_{\mu\nu}$ ($\Delta_{00} = 1$, $\Delta_{11} = \Delta_{22} = \Delta_{33} = -1$, $\Delta_{\mu\nu} = 0$, wenn $\mu \neq \nu$). Unter diesen Umständen betrachten wir die Welt als eine „Euklidische" mit dem metrischen Tensor $\Delta_{\mu\nu}$ und die $(\mu\rho\nu\sigma)$ als die Komponenten eines in dieser ebenen Welt eingebetteten Tensorfeldes vierter Stufe. Dabei spielen die $h_{\mu\nu}$ die Rolle der „Potentiale", deren Werte durch vier zusätzliche „Eichungsbedingungen"

$$[\alpha\alpha, \beta] = 0 \qquad (\beta = 0, 1, 2, 3) \tag{2}$$

fixiert werden können. (Hier, und im folgenden, bezieht sich die Summationsvorschrift nur auf griechische Indizes,

[1] Eine ausführlichere Fassung dieser Arbeit erscheint gleichzeitig im „Journal f. exp. und theor. Phys." (russisch).

und dabei in dem Sinne, dass $A_{\alpha\alpha}$ z. B. bedeutet $A_{00} - A_{11} - A_{22} - A_{33}$; dies erlaubt uns um den Unterschied zwischen ko- und kontravarianten Komponenten nicht mehr zu kümmern; $[\alpha\beta, \gamma]$ ist die gewöhnliche Bezeichnung des Christoffelschen Dreiindizessymbols.)

Die Gravitationsgleichungen im leeren Raume sind

$$(\mu\rho\nu\rho) = 0. \tag{3}$$

Diese sind unter den „Eichungsbedingungen" (2) den gewöhnlichen Wellengleichungen für Potentiale

$$\square h_{\mu\nu} = 0 \tag{4}$$

vollständig äquivalent.

Im folgenden betrachten wir ein quantenmechanisches kontinuierliches System, dessen klassische Bewegungsgleichungen in der Form (4) geschrieben werden können; unter den zusätzlichen Bedingungen (2), die, wie wir zeigen werden, mit der Schrödingergleichung des betrachteten quantenmechanischen Systems verträglich sind, wird dieses System dem Gravitationsfeld im leeren Raume identisch. Diese Behandlungsweise ist der bekannten Fermischen Methode der Quantisierung des elektromagnetischen Feldes gewissermassen analog: die Fermische Methode ist mit der Benutzung einer nicht-„eichinvarianten" Lagrangefunktion wesentlich verbunden, und so stützt sich auch unsere Methode der Quantisierung des Gravitationsfeldes auf die Benutzung von Grössen, die keine (sogar keine annähernden) Invarianten der allgemeinen Relativitätstheorie sind.

§ 2. Hamiltonsche Form und ebene Wellen.

Betrachten wir ein dynamisches Kontinuum mit zehn Feldgrössen $h_{\mu\nu}$ ($\mu \leqslant \nu$), die die Rolle der mechanischen Koordinaten des Systems spielen, und mit der Dichte der Lagrangeschen Funktion

$$[\alpha\alpha, \beta][\beta\gamma, \gamma] - [\alpha\beta, \gamma][\alpha\gamma, \beta] + \frac{1}{2}[\alpha\alpha, \beta][\gamma\gamma, \beta].$$

Ziemlich komplizierte Rechnungen, die wir hier der Kürze halber weglassen, zeigen, dass die entsprechende Dichte der

Hamiltonschen Funktion

$$
\begin{aligned}
&2\Big(p_{00}+\sum_l \frac{\partial h_{0l}}{\partial x_l}\Big)^2-\frac{1}{4}\Big(3p_{00}-\sum_l p_{ll}+2\sum_l \frac{\partial h_{0l}}{\partial x_l}\Big)^2-\\
&-\frac{1}{2}\sum_l\Big(-p_{0l}+\frac{\partial h_{00}}{\partial x_l}+\sum_m \frac{\partial h_{ml}}{\partial x_m}\Big)^2+\\
&+\frac{1}{4}\sum_n\Big(2p_{nn}+p_{00}-\sum_l p_{ll}+2\frac{\partial h_{0n}}{\partial x_n}\Big)^2+\\
&+\frac{1}{2}\sum_{l<m}\Big(p_{lm}+\frac{\partial h_{0m}}{\partial x_l}+\frac{\partial h_{0l}}{\partial x_m}\Big)^2+\frac{1}{8}\sum_l\Big(\frac{\partial h_{00}}{\partial x_l}\Big)^2-\\
&-\frac{1}{4}\sum_{lm}\Big(\frac{\partial h_{0l}}{\partial x_m}-\frac{\partial h_{0m}}{\partial x_l}\Big)^2+\frac{1}{4}\sum_{lm}\frac{\partial h_{00}}{\partial x_l}\frac{\partial h_{mm}}{\partial x_l}-\frac{1}{2}\Big(\sum_l \frac{\partial h_{0l}}{\partial x_l}\Big)^2-\\
&-\frac{1}{8}\sum_m\Big(\frac{\partial}{\partial x_m}\sum_l h_{ll}\Big)^2+\frac{1}{4}\sum_{lmn}\Big(\frac{\partial h_{ln}}{\partial x_m}\Big)^2+\\
&+\frac{1}{2}\sum_{lmn}\Big(\frac{\partial h_{lm}}{\partial x_m}\frac{\partial h_{ln}}{\partial x_n}-\frac{\partial h_{lm}}{\partial x_n}\frac{\partial h_{ln}}{\partial x_m}\Big)
\end{aligned}
$$

ist, wo $p_{\alpha\beta}$ die mit den mechanischen Koordinaten $h_{\alpha\beta}$ konjugierten Impulse sind, und dass die entsprechenden Bewegungsgleichungen (4) sind. Hier, wie auch im Folgenden, nehmen lateinische Indizes nur die Werte 1, 2, 3 an. Wir haben auch, der Einfachheit halber, die Lichtgeschwindigkeit gleich 1 gesetzt, und die Newtonsche Gravitationskonstante gleich $1/16\pi$. (Den Wert der Gravitationskonstante kann der Leser leicht prüfen, wenn er unsere Formeln z. B. mit den Formeln 58.1 und 59.4 des Eddingtonschen Buches „Relativitätstheorie in mathematischer Behandlung", Berlin, Springer 1925, vergleicht).

Führen wir nun die Fourierentwicklung

$$h_{\alpha\beta}=\frac{1}{(2\pi)^{3/2}}\int d\mathfrak{k}\,[h_{\alpha\beta,\mathfrak{k}}\,e^{-i(\omega t-\mathfrak{k}\mathfrak{r})}+h^{+}_{\alpha\beta,\mathfrak{k}}\,e^{i(\omega t-\mathfrak{k}\mathfrak{r})}]$$

(wo $\omega=|\mathfrak{k}|$) ein. Die Rechnungen zeigen, dass die Hamiltonsche Funktion des Feldes dabei in der Form

$$H = \int d\mathfrak{k}\,\omega^2 \Big\{ \tfrac{1}{2}\big(h^+_{00,\mathfrak{k}} + \sum_l h^+_{ll,\mathfrak{k}}\big)\big(h_{00,\mathfrak{k}} + \sum_l h_{ll,\mathfrak{k}}\big) +$$

$$+ \sum_{l \neq m} (h_{lm,\mathfrak{k}} h^+_{lm,\mathfrak{k}} - h^+_{ll,\mathfrak{k}} h_{mm,\mathfrak{k}}) - 2 \sum_l h_{0l,\mathfrak{k}} h^+_{0l,\mathfrak{k}} \Big\} \tag{5}$$

geschrieben werden kann. (Die Reihenfolge der Faktoren scheint zunächst willkürlich; es ist aber aus dem folgenden klar, dass diese Reihenfolge die einzige ist, die bei der Übertragung auf das Quantengebiet die Vermeidung der „Nullpunktsenergie" gestattet.)

Die Bedingungen (2) nehmen nach der Fourierentwicklung die folgende Form an:

$$\left.\begin{aligned} &\tfrac{1}{2}\omega(h_{00,\mathfrak{k}} + \sum_l h_{ll,\mathfrak{k}}) + \sum_l \mathfrak{k}_l h_{0l,\mathfrak{k}} = 0, \\ &\omega h_{0l,\mathfrak{k}} + \sum_m \mathfrak{k}_m h_{ml,\mathfrak{k}} + \tfrac{1}{2}\mathfrak{k}_l (h_{00,\mathfrak{k}} - \sum_m h_{mm,\mathfrak{k}}) = 0 \\ &(l = 1, 2, 3). \end{aligned}\right\} \tag{6}$$

Die Zahl der unabhängigen $h_{\mu\nu,\mathfrak{k}}$ (bei gegebenem $\mathfrak{k}$) ist also $10 - 4 = 6$. Man kann aber leicht zeigen, dass bei vielen Fragen, wo es sich z. B. um die Energieübertragung durch die Gravitationswellen handelt, die Zahl der unabhängigen Polarisationen noch kleiner als 6 ist. Betrachten wir z. B. den Fall $\mathfrak{k} \parallel z$. Die Bedingungen (6) werden dann zu

$$h_{11,\mathfrak{k}} + h_{22,\mathfrak{k}} = h_{00,\mathfrak{k}} + 2h_{03,\mathfrak{k}} + h_{33,\mathfrak{k}} = h_{01,\mathfrak{k}} + h_{31,\mathfrak{k}} =$$

$$= h_{02,\mathfrak{k}} + h_{32,\mathfrak{k}} = 0. \tag{6'}$$

Der Energiegehalt der Gravitationswelle kann, mit Berücksichtigung von (6′), in der Form

$$2d\mathfrak{k}\,\omega^2 (h^+_{12,\mathfrak{k}} h_{12,\mathfrak{k}} + h^+_{11,\mathfrak{k}} h_{11,\mathfrak{k}})$$

geschrieben werden. (In solche Form kann man den Integrand von (5) mit Hilfe der Bedingungen (6′) leicht bringen, wenn alle $h_{\mu\nu,\mathfrak{k}}$ und $h^+_{\mu\nu,\mathfrak{k}}$ kommutativ sind; später werden wir aber sehen, dass auch im Quantengebiet, wo die h und die h^+ nicht mehr kommutativ sind, ein ähnliches Resultat gilt.) Wir können also ohne den Energiegehalt der Gravitationswelle zu ändern, die 10 Amplituden $h_{\mu\nu,\mathfrak{k}}$ ausser den Bedingungen (6) noch vier weiteren Bedingungen unterwerfen.

Z. B. können wir (auch wenn $\mathfrak{k}$ nicht $\parallel z$ ist) diese weiteren Bedingungen in der folgenden (zwar relativistisch nicht invarianten) Form wählen:

$$h_{00,\mathfrak{k}} = h_{01,\mathfrak{k}} = h_{02,\mathfrak{k}} = h_{03,\mathfrak{k}} = 0.$$

Wir sehen also u. a., dass für die Energieübertragung nur die transversalen Gravitationswellen, und zwar mit zwei unabhängigen Polarisationen, wesentlich sind. Die Sache steht aber ganz anders, wenn es sich nicht um die Energieübertragung, sondern z. B. um die Gravitationszusammenwirkung von zwei schweren Körpern handelt: unten werden wir sehen, dass für diese Zusammenwirkung gerade die longitudinalen $h_{\mu\nu}$-Wellen von allergrösster Bedeutung sind.

§ 3. Vertauschungsrelationen und Eigenwerte der Energie.

Die Heisenberg-Paulischen V.-R. lauten [1]

$$\left.\begin{aligned}
&[h_{\alpha\beta}(\mathfrak{r}),\, h_{\alpha'\beta'}(\mathfrak{r}')] = 0,\\
&[p_{\alpha\beta}(\mathfrak{r}),\, p_{\alpha'\beta'}(\mathfrak{r}')] = 0,\\
&[p_{\alpha\beta}(\mathfrak{r}),\, h_{\alpha'\beta'}(\mathfrak{r}')] = \frac{h}{i}\,\delta_{\alpha\alpha'}\delta_{\beta\beta'}\delta(\mathfrak{r}-\mathfrak{r}') \qquad (\alpha \leqq \beta,\ \alpha' \leqq \beta').
\end{aligned}\right\} \quad (7)$$

Nach der Fourierentwicklung bekommen wir

$$\left.\begin{aligned}
&[h_{\alpha\beta,\mathfrak{k}},\, h_{\alpha'\beta',\mathfrak{k}'}] = 0, \quad [h^+_{\alpha\beta,\mathfrak{k}},\, h^+_{\alpha'\beta',\mathfrak{k}'}] = 0,\\
&[h^+_{00,\mathfrak{k}},\, h_{00,\mathfrak{k}'}] = [h^+_{00,\mathfrak{k}},\, h_{ll,\mathfrak{k}'}] = [h^+_{ll,\mathfrak{k}},\, h_{ll,\mathfrak{k}'}] = -\frac{h}{2\omega}\,\delta(\mathfrak{k}-\mathfrak{k}'),\\
&[h^+_{00,\mathfrak{k}},\, h_{0l,\mathfrak{k}'}] = [h^+_{ll,\mathfrak{k}},\, h_{0m,\mathfrak{k}'}] = [h^+_{0l,\mathfrak{k}},\, h_{mn,\mathfrak{k}'}] = 0,\\
&[h^+_{00,\mathfrak{k}},\, h_{lm,\mathfrak{k}'}] = [h^+_{nn,\mathfrak{k}},\, h_{lm,\mathfrak{k}'}] = 0 \quad (l \neq m),\\
&[h^+_{ll,\mathfrak{k}},\, h_{mm,\mathfrak{k}'}] = \frac{h}{2\omega}\,\delta(\mathfrak{k}-\mathfrak{k}') \quad (l \neq m),\\
&[h^+_{0l,\mathfrak{k}},\, h_{0m,\mathfrak{k}'}] = \frac{h}{2\omega}\,\delta_{lm}\delta(\mathfrak{k}-\mathfrak{k}'),\\
&[h^+_{lm,\mathfrak{k}},\, h_{pq,\mathfrak{k}'}] = -\frac{h}{2\omega}\,\delta_{lp}\delta_{mq}\delta(\mathfrak{k}-\mathfrak{k}') \quad (l < m,\ p < q).
\end{aligned}\right\} \quad (8)$$

---

[1] h bedeutet hier $h/2\pi$.

Wir führen die Operatoren

$$A = \tfrac{1}{2}\,\omega\,(h_{00,\mathfrak{k}} + \sum_l h_{ll,\mathfrak{k}}) + \sum_l \mathfrak{k}_l h_{0l,\mathfrak{k}},$$

$$B_l = \omega h_{0l,\mathfrak{k}} + \sum_m \mathfrak{k}_m h_{ml,\mathfrak{k}} + \tfrac{1}{2}\,\mathfrak{k}_l\,(h_{00,\mathfrak{k}} - \sum_m h_{mm,\mathfrak{k}}),$$

ein. Die Rechnung zeigt, dass infolge von (8) alle acht Operatoren A, B_l, A^+, B_l^+ ($l = 1, 2, 3$) untereinander vertauschbar sind. Mit der Hamiltonschen Funktion (5) sind sie aber nicht vertauschbar, d. h. sie sind keine Bewegungsintegrale. Es lässt sich jedoch leicht zeigen, dass die Poissonklammern eines jeden dieser Operatoren und der Hamiltonschen Funktion, d. h. die Veränderungsgeschwindigkeiten dieser Operatoren, im speziellen Fall, wo

$$A = B_l = A^+ = B_l^+ = 0,$$

alle verschwinden. Das bedeutet, dass die Bedingungen (6) (und die entsprechenden adjungierten) miteinander und mit den quantenmechanischen Bewegungsgleichungen verträglich sind.

Die Vertauschungsrelationen (8), der Hamiltonsche Operator (5) und die Zusatzbedingungen (6) (zusammen mit dem später einzuführenden Ansatz für die Wechselwirkung zwischen Gravitationsfeld und Materie) bilden die Grundlage der hier vorgeschlagenen Quantentheorie der Gravitation. Es sei bemerkt, dass die Aufstellung eines quantenmechanischen Hamiltonschen Operators mit Hilfe des Korrespondenzprinzips niemals eindeutig sein kann: es ist immer möglich, die Hamiltonsche Funktion durch Hinzufügung der mit $h \to 0$ verschwindenden Zusatzglieder abzuändern (z. B. des „Spinglieds" in der Theorie des Elektrons); sogar die Relativitätsforderungen reichen im allgemeinen nicht aus, um diese „Spinglieder" eindeutig zu fixieren. Doch glauben wir, dass in der Gravitationstheorie die Hinzufügung solcher Zusatzglieder unnötig ist.

Nun zur Berechnung der Eigenwerte der Energie! Das geschieht wie in der Quantenelektrodynamik durch die Einführung der neuen Variablen ξ, die den V. R.

$$[\xi, \xi^+] = 1$$

genügen. Dann sind bekanntlich die Eigenwerte von $\xi\xi^+$ gleich $n+1$, und die von $\xi^+\xi$ gleich n, wo n eine positive ganze Zahl oder Null ist.

Hier gelingt es nicht, solche ξ - Variablen auf symmetrische Weise einzuführen. Eine mögliche Lösung des Problems sieht so aus

$$\frac{1}{2}\left(h_{00,\mathfrak{k}}+\sum_l h_{ll,\mathfrak{k}}\right)=\sqrt{\frac{h}{2\omega d\mathfrak{k}}}\,\xi_{00,\mathfrak{k}}\,e^{i\omega t},$$

$$h_{11,\mathfrak{k}}=\sqrt{\frac{h}{2\omega d\mathfrak{k}}}\left(\frac{\xi^+_{11,\mathfrak{k}}}{\sqrt{3}}\,e^{-i\omega t}+\frac{\xi_{22,\mathfrak{k}}}{\sqrt{3}}\,e^{i\omega t}+\xi_{33,\mathfrak{k}}\,e^{i\omega t}\right),$$

$$h_{22,\mathfrak{k}}=\sqrt{\frac{h}{2\omega d\mathfrak{k}}}\left(\frac{\xi^+_{11,\mathfrak{k}}}{\sqrt{3}}\,e^{-i\omega t}+\frac{\xi_{22,\mathfrak{k}}}{\sqrt{3}}\,e^{i\omega t}-\xi_{33,\mathfrak{k}}\,e^{i\omega t}\right),$$

$$h_{33,\mathfrak{k}}=\sqrt{\frac{h'}{2\omega d\mathfrak{k}}}\left(\frac{\xi^+_{11,\mathfrak{k}}}{\sqrt{3}}\,e^{-i\omega t}-\frac{2}{\sqrt{3}}\,\xi_{22,\mathfrak{k}}\,e^{i\omega t}\right),$$

$$h_{lm,\mathfrak{k}}=\sqrt{\frac{h}{2\omega d\mathfrak{k}}}\,\xi_{lm,\mathfrak{k}}\,e^{i\omega t},\quad (l\neq m),$$

$$h_{0l,\mathfrak{k}}=\sqrt{\frac{h}{2\omega d\mathfrak{k}}}\,\xi^+_{0l,\mathfrak{k}}\,e^{-i\omega t}.$$

Die Hamiltonsche Funktion wird in den neuen Variablen zu

$$\left.\begin{aligned}H=\sum_k h\omega\,(\xi^+_{00,\mathfrak{k}}\xi_{00,\mathfrak{k}}+\xi_{12,\mathfrak{k}}\xi^+_{12,\mathfrak{k}}+\xi_{23,\mathfrak{k}}\xi^+_{23,\mathfrak{k}}+\xi_{13,\mathfrak{k}}\xi^+_{13,\mathfrak{k}}-{}\\-\xi^+_{01,\mathfrak{k}}\xi_{01,\mathfrak{k}}-\xi^+_{02,\mathfrak{k}}\xi_{02,\mathfrak{k}}-\xi^+_{03,\mathfrak{k}}\xi_{03,\mathfrak{k}}-\xi_{11,\mathfrak{k}}\xi^+_{11,\mathfrak{k}}+{}\\+\xi^+_{22,\mathfrak{k}}\xi_{22,\mathfrak{k}}+\xi^+_{33,\mathfrak{k}}\xi_{33,\mathfrak{k}}).\end{aligned}\right\}\quad(5')$$

Darum sind die Eigenwerte der Energie (für jeden Wert von $\mathfrak{k}$)

$$h\omega\,(n_{00}+n_{12}+n_{23}+n_{31}-n_{01}-n_{02}-n_{03}-n_{11}+n_{22}+n_{33}+2),$$

wo n_{00}, $n_{12}, \ldots$ zehn Quantenzahlen sind ($n=0, 1, 2, \ldots$).

Die Bedingungen (6) machen diesen Ausdruck positiv-definit. Um das einzusehen, betrachten wir wieder den Fall $\mathfrak{k} \parallel z$. Aus (6′) erhalten wir dann folgende Bedingungen:

$$\xi_{00,\mathfrak{k}}\,e^{i\omega t}+\xi^+_{03,\mathfrak{k}}\,e^{-i\omega t}=0,\qquad \xi^+_{01,\mathfrak{k}}\,e^{-i\omega t}+\xi_{13,\mathfrak{k}}\,e^{i\omega t}=0,$$

$$\xi^+_{02,\mathfrak{k}}\,e^{-i\omega t}+\xi_{23,\mathfrak{k}}\,e^{i\omega t}=0,\qquad \xi^+_{11,\mathfrak{k}}\,e^{-i\omega t}+\xi_{22,\mathfrak{k}}\,e^{i\omega t}=0.$$

Es folgt daraus, dass

$$n_{01} = n_{31} + 1, \quad n_{02} = n_{23} + 1, \quad n_{22} = n_{11} + 1, \quad n_{03} = n_{00} + 1.$$

Die Eigenwerte der Energie für diesen $\mathfrak{k}$ werden zu

$$h\omega\,(\xi^{+}_{12,\mathfrak{k}}\,\xi_{12,\mathfrak{k}} + \xi^{+}_{33,\mathfrak{k}}\,\xi_{33,\mathfrak{k}}) = h\omega\,(n_{12} + n_{33}).$$

Wir sehen also, dass die Energie des Gravitationsfeldes aus den positiven Gravitationsquanten besteht, und zwar von je zwei Polarisationen für jeden Wellenvektor $\mathfrak{k}$. Analog zu dem klassischen Fall spielen auch hier nur die transversalen Gravitationsschwingungen eine Rolle: z. B. für $\mathfrak{k} \parallel z$ eine h_{12}- und eine $\frac{1}{2}(h_{11} - h_{22})$-Schwingung.

Dabei entstehen keine „Nullpunktsenergieglieder": das ist infolge einer zweckmässig gewählten Reihenfolge der Faktoren im Ausdruck (5) gelungen.

§ 4. Ein wenig Gedankenexperimentieren!

Um den physikalischen Inhalt der Quantentheorie des Gravitationsfeldes etwas näher zu verstehen, betrachten wir die Messung einer der hier vorkommenden Feldgrössen, z. B. des Christoffelschen Dreiindizessymbols [00,1]. Die klassischen Bewegungsgleichungen Einsteins lauten in unserem Fall (alle $h_{\mu\nu} \ll 1$):

$$\frac{d^2x}{dt^2} = \frac{\partial h_{01}}{\partial t} - \frac{1}{2}\frac{\partial h_{00}}{\partial x} = [00,1]. \tag{9}$$

Nach dem Vorbilde von Bohr und Rosenfeld[1] betrachten wir die Messung eines raumzeitlichen Mittelwertes von [00,1] im Volumen V und in dem Zeitintervall T. Nehmen wir einen Probekörper vom Volumen V. Seine Masse sei ρV. Die obere Bewegungsgleichung, die nur dann gilt, wenn die Geschwindigkeit des Probekörpers klein gegenüber der Lichtgeschwindigkeit ist, macht die folgende Messung möglich: messen wir den Impuls des Probekörpers am Anfang und dann am Ende eines Zeitintervalls T, so ist der gesuchte

[1] N. Bohr und L. Rosenfeld, Dansk. Vidensk. Selskab., Math.-fys. Meddel. 12, 8, 1933.

Mittelwert definitionweise gleich

$$\frac{(p_x)_{t+T}-(p_x)_t}{\rho VT}.$$

Die Messung von [00,1] ist daher mit einer Unbestimmtheit verbunden von der Grössenordnung

$$\Delta\,[00,1] \approx \Delta p_x/\rho VT, \tag{10}$$

wo Δp_x die Unbestimmtheit des Impulses ist. Die Dauer der Impulsmessung sei Δt (selbstverständlich ist $\Delta t \ll T$); Δx sei die mit der Impulsmessung verbundene Unbestimmtheit der Koordinate. Die Unbestimmtheit Δp_x besteht aus zwei Gliedern: aus dem gewöhnlichen quantenmechanischen Glied $h/\Delta x$ und aus einem Glied, das mit dem Gravitationsfeld verknüpft ist, das von dem Probekörper selbst wegen seines Rückstosses bei der Messung erzeugt wird. Wegen der Einsteinschen Gravitationsgleichung $\Box h_{01} = \rho v_x$ muss die Unbestimmtheit von h_{01}, die infolge der unbestimmten Rückstossgeschwindigkeit $\Delta x/\Delta t$ entsteht, von der Grössenordnung $\rho\,\frac{\Delta x}{\Delta t}\cdot\Delta t^2$ sein. Es ist aus (9) ersichtlich, dass die entsprechende Unbestimmtheit von [00,1] von der Grössenordnung $\rho\Delta x$ ist, und damit während jeder Impulsmessung eine mit dem Gravitationsfeld verbundene zusätzliche Impulsunbestimmtheit von der Grössenordnung $\rho\Delta x\cdot\rho V\Delta t$ entsteht. Um den Vergleich mit den gewöhnlichen Messeinheiten zu erleichtern, sehen wir hier (bis zum Ende dieses Paragraphs) von unserer Vorschrift $c = 1$, $G = 1/16\pi$ ab. Wir erhalten für den Impuls

$$\Delta p_x \approx \frac{h}{\Delta x} + G\rho^2 V\Delta x\Delta t.$$

Es kann gezeigt werden (analog zu den Überlegungen von Bohr und Rosenfeld), dass das zweite Glied gegenüber dem ersten beliebig klein gemacht werden kann. Für die beste Ausführung der [00,1]-Messung scheint es aber zweckmässiger, Δp_x zum Minimum zu machen, d. h. für die beiden Glieder die gleiche Grössenordnung zu schaffen. Dazu soll Δx von der Grössenordnung

$$\Delta x \approx \frac{1}{\rho}\left(\frac{h}{G V \Delta t}\right)^{1/3}$$

gewählt werden. Für $\Delta[00,1]$ erhalten wir

$$\Delta\,[00,1] \gtrsim \frac{h^{1/2} G^{1/2} \Delta t^{1/2}}{V^{1/2} T}. \qquad (11)$$

Eine absolut genaue Messung des Schwerefeldes wäre also dann möglich, wenn eine beliebig schnelle Impulsmessung möglich wäre. Zwei Umstände machen es aber unmöglich, eine beliebig schnelle Impulsmessung auszuführen: erstens soll nach der Definition der Messung $\Delta x \ll V^{1/3}$ sein, und das führt zu

$$\Delta t \gg \frac{h}{\rho^2 G V^{5/3}}.$$

Zweitens kann nach der Relativitätstheorie niemals Δx grösser als $c\Delta t$ werden; und das führt zu

$$\Delta t \gtrsim \frac{h^{1/3}}{c^{2/3} \rho^{2/3} V^{1/3} G^{1/3}}.$$

Es folgt nach (11), dass niemals $\Delta\,[00,1]$ kleiner als

$$\frac{h}{\rho T V^{4/3}} \quad \text{oder} \quad \frac{h^{2/3} G^{1/3}}{c^{1/3} \rho^{1/3} V^{2/3} T}$$

gemacht werden kann. Von diesen beiden Grenzen ist für den Fall eines leichten Probekörpers ($\rho V \lesssim h^{1/2} c^{1/2} / G^{1/2}$, d. h. kleiner als etwa 0,01 mgr) die erste die einzige wesentliche. Für einen schwereren Probekörper ist die zweite die wesentlichste. Es ist klar, dass für eine möglichst genaue [00,1]-Messung gerade schwere Probekörper zu empfehlen sind, und das bedeutet, dass nur die zweite Grenze theoretisch wichtig ist. Wir haben schliesslich

$$\Delta\,[00,1] \gtrsim \frac{h^{2/3} G^{1/3}}{c^{1/3} \rho^{1/3} V^{2/3} T}. \qquad (12)$$

Es ist also klar, dass im Gebiet, wo alle $h_{\mu\nu}$ klein gegen 1 sind (das ist gerade die Bedeutung des Wortes „schwach" im Titel dieser Arbeit), die Genauigkeit der Schweremessungen beliebig hoch gesteigert werden kann: denn in diesem Erscheinungsgebiet gelten die angenäherten linearisierten Gleichungen (1), es gilt also auch das Superpositionsprinzip, und es ist daher immer möglich, einen Probekörper mit beliebig hohem ρ zu beschaffen. Daraus schliessen wir, dass es möglich ist, wie es z. B. diese Arbeit zu tun versucht, im Rahmen der speziellen Relativitätstheorie (d. h. wenn das raumzeitliche Kontinuum ein „Euklidisches" ist) eine durchaus konsequente Quantentheorie der Gravitation aufzubauen. Im Gebiet der allgemeinen Relativitätstheorie, wo die Abweichungen von der „Euklidizität" beliebig gross sein können, steht aber die Sache ganz anders. Denn der Gravitationsradius des zur Messung dienenden Probekörpers ($G\rho V/c^2$) soll keineswegs grösser als seine linearen Abmessungen ($V^{1/3}$) sein; daraus entsteht eine obere Grenze für seine Dichte ($\rho \lesssim c^2/GV^{2/3}$). Die Messungsmöglichkeiten sind also in diesem Gebiet noch mehr beschränkt, als es sich aus den quantenmechanischen V.-R. schliessen lässt. Ohne eine tiefgehende Umarbeitung der klassischen Begriffe scheint es daher wohl kaum möglich, die Quantentheorie der Gravitation auch auf dieses Gebiet auszudehnen.

§ 5. Wechselwirkung mit der Materie.

Ein korrespondenzmässig richtiger Ansatz für die Energie der Wechselwirkung zwischen Gravitationsfeld und Materie kann aus der von V. Fock[1] aufgestellten allgemeinrelativistischen Form der Diracschen Wellengleichung gewonnen werden. Wenn alle $h_{\mu\nu}$ klein gegen 1 sind, kann man diese Gleichung für den Fall des verschwindenden elektromagnetischen Feldes in der folgenden Form schreiben:

$$\frac{h}{i}\sum_{k=0}^{3} e_k\alpha_k\left(\frac{\partial}{\partial x_k}-\frac{1}{2}\sum_{l=0}^{3} e_l h_{kl}\frac{\partial}{\partial x_l}\right)\psi+$$

$$+\left(\frac{1}{8}\frac{h}{i}\sum_{k=0}^{3}\frac{\partial h_{00}}{\partial x_k}\alpha_k-m\beta\right)\psi=0,$$

[1] V. Fock, ZS. f. Phys. 57, 261, 1929.

wo

$$e_0 = 1, \quad e_1 = e_2 = e_3 = -1$$

und

$$\alpha_0 = 1, \quad \alpha_1 = \rho_1\sigma_1, \quad \alpha_2 = \rho_1\sigma_2, \quad \alpha_3 = \rho_1\sigma_3, \quad \beta = \rho_3$$

(Diracsche Matrizen). Führen wir statt der vierkomponentigen ψ-Funktion zwei zweikomponentige Funktionen χ und φ ein

$$\psi_1 = \chi_1 e^{-imt/h}, \quad \psi_2 = \chi_2 e^{-imt/h}, \quad \psi_3 = \varphi_1 e^{-imt/h}, \quad \psi_4 = \varphi_2 e^{-imt/h},$$

so wird bei abnehmender Geschwindigkeit des Teilchens χ gleich Null, und φ zu seiner nichtrelativistischen Wellenfunktion. Es lässt sich aus der Schrödingergleichung für diese φ_x ersehen, dass die Energie der Wechselwirkung zwischen dem Teilchen und dem Gravitationsfeld die Form

$$V = \frac{m}{2} h_{00} + \frac{h}{i} \sum_k h_{0k} \frac{\partial}{\partial x_k} - \frac{h^2}{2m} \sum_{kl} h_{kl} \frac{\partial^2}{\partial x_k \partial x_l} - \frac{h^2}{4m} \sum_{kl} \frac{\partial h_{kl}}{\partial x_l} \frac{\partial}{\partial x_k} +$$
$$+ \frac{h}{4i} \sum_l \frac{\partial h_{0l}}{\partial x_l} + \frac{h^2}{4mi} \sum_{jklm} \sigma_j e_{jlm} \frac{\partial h_{km}}{\partial x_l} \frac{\partial}{\partial x_k} + \frac{h}{4} \sum_{jlm} \sigma_j e_{jlm} \frac{\partial h_{0m}}{\partial x_l}$$

hat, wo e_{jlm} der schiefsymmetrische Einheitstensor ist (d. h. $e_{123} = 1$ und e_{jlm} ist in Bezug auf jedes Paar seiner Indizes antisymmetrisch). Wenn die Wellenlänge der Gravitationsstörungen genügend gross ist, vereinfacht sich dieser Ausdruck zu

$$V = \frac{m}{2} h_{00} + \frac{h}{i} \sum_k h_{0k} \frac{\partial}{\partial x_k} - \frac{h^2}{2m} \sum_{kl} h_{kl} \frac{\partial^2}{\partial x_k \partial x_l}. \tag{13}$$

Den Ansatz (13) werden wir im folgenden benutzen. Es sei bemerkt, dass sogar die einfachen korrespondenzmässigen Betrachtungen, ohne den Umweg über die Dirac-Focksche Gleichung, auch zum Ansatz (13) für die Wechselwirkungsenergie führen.

§ 6. Energieübertragung durch die Gravitationswellen.

Eine der einfachsten Anwendungen der oben skizzierten Quantentheorie der Gravitation besteht in der Berechnung der Energieausstrahlung in Form der von materiellen Systemen emittierten Gravitationswellen. Hier sollen wir von den Bedingungen (6) oder (6′) zusammen mit $h_{00,\mathfrak{k}} = h_{01,\mathfrak{k}} = h_{02,\mathfrak{k}} = h_{03,\mathfrak{k}} = 0$ Gebrauch machen. In den ξ-Variablen (siehe oben, § 3) liefert das für den Fall $\mathfrak{k} \parallel z$

$$V = \frac{m}{8\pi}\sqrt{\frac{h\,d\mathfrak{k}}{\pi\omega}}\left\{[\xi(\dot{x}_1^2 - \dot{x}_2^2) + 2\eta\,\dot{x}_1\dot{x}_2]\,e^{i\mathfrak{k}\mathfrak{r}} + \right.$$
$$\left. + [\xi^+(\dot{x}_1^2 - \dot{x}_2^2) + 2\eta^+\dot{x}_1\dot{x}_2]\,e^{-i\mathfrak{k}\mathfrak{r}}\right\}, \qquad (14)$$

wo der Kürze halber $\dot{x}_k$ statt $\frac{h}{mi}\frac{\partial}{\partial x_k}$ geschrieben ist, und ξ und η statt ξ_{33} und ξ_{12}. Nennen wir den Anfangs- bzw. den Endzustand des emittierenden Teilchens k bzw. l, den Anfangszustand bzw. den Endzustand der Gravitationseigenschwingung k' bzw. l'. Die Quantenmechanik liefert für die Wahrscheinlichkeit des Überganges pro Zeiteinheit den bekannten Ausdruck

$$\frac{2\pi}{h}\,\delta(E_l + E_{l'} - E_k - E_{k'})\,|(kk'\,|\,V\,|\,ll')|^2.$$

Mit Benutzung der bekannten Werte der Oszillatorenmatrixelemente schliessen wir daraus, dass die Wahrscheinlichkeit der spontanen Ausstrahlung eines Gravitationsquants pro Zeiteinheit mit einem gegebenen $\mathfrak{k}$ ($\mathfrak{k} \parallel z$) und mit der ξ-Polarisation

$$\frac{m^2}{32\pi^2}\frac{d\mathfrak{k}}{\omega}\,|(k\,|\,\dot{x}_1^2 - \dot{x}_2^2\,|\,l)|^2\,\delta(E_l - E_k + h\omega)$$

ist, und mit der η-Polarisation:

$$\frac{m^2}{32\pi^2}\frac{d\mathfrak{k}}{\omega}\,|(k\,|\,2\dot{x}_1\dot{x}_2\,|\,l)|^2\,\delta(E_l - E_k + h\omega).$$

Die Wahrscheinlichkeit des Überganges $k \to l$ mit der gleichzeitigen Emission eines Gravitationsquantes in die Öffnung eines Kegels $d\Omega$ ($\parallel z$) pro Zeiteinheit ist daher

$$\frac{d\Omega}{8\pi^2}\frac{m^2\omega}{h}\left\{\left|\left(k\left|\frac{\dot{x}_1^2 - \dot{x}_2^2}{2}\right|l\right)\right|^2 + |(k\,|\,\dot{x}_1\dot{x}_2\,|\,l)|^2\right\}.$$

Es ist nicht schwierig, diesen Ausdruck für eine beliebige Richtung (nicht nur $\parallel z$) zu verallgemeinern, und dann über alle Richtungen zu integrieren. Die Rechnung führt zu dem folgenden Ergebnis: die totale Wahrscheinlichkeit (pro Zeiteinheit) des Überganges $k \to l$ mit gleichzeitiger Emission eines Gravitationsquantes von der Frequenz $\omega = (E_k - E_l)/h$ ist

$$\frac{m^2\omega}{10\pi h}\left\{\sum_{pq} |(k\,|\,\dot{x}_p\dot{x}_q\,|\,l)|^2 - \tfrac{1}{3}\,|(k\,|\sum_p \dot{x}_p^2\,|\,l\,|^2\right\}.$$

Es ist nicht schwierig, diesen Ausdruck auch auf ein beliebiges System materieller Teilchen zu verallgemeinern. Für die Energie, die ein solches System in der Form der beim Übergang $k \to l$ pro Zeiteinheit emittierten Gravitationswellen verliert, erhalten wir einen Ausdruck

$$\frac{\omega^2}{10\pi}\left\{\sum_{pq} |(k\,|\sum m\dot{x}_p\dot{x}_q\,|\,l)|^2 - \tfrac{1}{3}\,|(k\,|\sum_p\sum m\dot{x}_p^2\,|\,l)\,|^2\right\} \tag{15}$$

(das Summenzeichen Σ ohne Indizes bedeutet die Summation über verschiedene Teilchen). Diese Formel ist eine quantentheoretische Verallgemeinerung des bekannten Einsteinschen Resultats.

In der Tat lautet der Einsteinsche Ausdruck für die in der Form der Gravitationswellen pro Zeiteinheit ausgestrahlte Energie (bei $G = 1/16\pi$)

$$\frac{1}{80\pi}\left\{\sum_{pq}\left(\frac{d^3}{dt^3}\sum m x_p x_q\right)^2 - \frac{1}{3}\left(\sum_p \frac{d^3}{dt^3}\sum m x_p^2\right)^2\right\}.\text{[1]}$$

Wenn $I_{pq} \equiv \sum m x_p x_q$ durch eine Fourierreihe von der Form

$$I_{pq} = \sum_{k=-\infty}^{+\infty} I_{pq}^{(k)}\, e^{ik\omega_0 t}$$

darstellbar ist, so wird dieser klassische Ausdruck für die Ausstrahlung der Energie in der Frequenz $\omega = k\omega_0$ zu

$$\frac{\omega^6}{40\pi}\left\{\sum_{pq} |I_{pq}^{(k)}|^2 - \tfrac{1}{3}\left|\sum_p I_{pp}^{(k)}\right|^2\right\}. \tag{16}$$

[1] Bei Einstein selbst (Berl. Ber. 1918, S. 154) steht infolge eines Rechenfehlers $1/160\pi$ statt $1/80\pi$. Die späteren Rechnungen Eddingtons (vgl. sein Lehrbuch oder Proc. Roy. Soc. **102**, 281, 1922) führen zu dem richtigen Koeffizienten.

Auf der anderen Seite haben wir

$$(k\,|\sum m\dot{x}_p\dot{x}_q\,|\,l) = \sum\sum_j (k\,|\sqrt{m}\dot{x}_p\,|\,j)\,(j\,|\sqrt{m}\,\dot{x}_q\,|\,l) =$$

$$= -\sum\sum_j m\,(k\,|\,x_p\,|\,j)\,(j\,|\,x_q\,|\,l)\,\omega_{kj}\omega_{jl}.$$

Im Gebiet niedriger Frequenzen und hoher Quantenzahlen können wir, wie man es z. B. für den Fall eines Rotators leicht berechnet, ω^2 statt $2\omega_{kj}\omega_{jl}$ setzen (ω = der ausgestrahlten Frequenz), und das liefert näherungsweise

$$(k\,|\sum m\dot{x}_p\dot{x}_q\,|\,l) = -\frac{\omega^2}{2}\,(k\,|\,I_{pq}\,|\,l).$$

Dann wird der Ausdruck (15) zu

$$\frac{\omega^6}{40\pi}\left\{\sum_{pq}|\,(k\,|\,I_{pq}\,|\,l)\,|^2 - \tfrac{1}{3}\,|\sum_p (k\,|\,I_{pp}\,|\,l)\,|^2\right\}.$$

Wenn wir statt der Matrixelemente Fourieramplituden schreiben, geht dieser letzte Ausdruck in den klassischen Einsteinschen (16) über. Im Grenzfall $h \to 0$ liefert also die Quantentheorie der Gravitation dieselben Resultate wie die klassische Theorie Einsteins.

§ 7. Herleitung des Newtonschen Gravitationsgesetzes.

Dirac hat gezeigt,[1] dass allerlei Wechselwirkungen zwischen Ladung und Ladung immer als durch Vermittelung eines intermediären Agens, nämlich des quantisierten Feldes, realisiert interpretiert werden können. Wir werden hier zeigen, dass eine ähnliche Situation auch im Gebiet der Gravitationserscheinungen vorliegt. Zunächst klingt es etwas paradox, denn die beiden Ausdrücke der Wechselwirkung zwischen Feld und Materie sind fast genau gleich [$e\Phi$ ist das Hauptglied der Wechselwirkung im elektromagnetischen Fall; wenn wir statt der Ladung e $m/2$ schreiben und das skalare Potential Φ durch das skalare Potential h_{00} ersetzen, erhalten wir das Hauptglied von (13)], und dennoch soll ein und das-

[1] P. A. M. Dirac, Proc. Roy. Soc. 136, 453, 1932.

selbe Schema in einem Fall die Abstossung von Partikeln gleicher Art (Coulombkräfte), im anderen Fall aber die Anziehung (Newtonkräfte) erklären. Die Lösung des Paradoxons besteht darin, dass in der Quantenelektrodynamik die V.-R. für das Potential Φ

$$[\Phi^{+}_{\mathfrak{k}}, \Phi_{\mathfrak{k}'}] = \frac{h}{2\omega}\,\delta(\mathfrak{k}-\mathfrak{k}')$$

lauten, indem bei uns eine andere V.-R., nämlich

$$[h^{+}_{00,\mathfrak{k}}, h_{00,\mathfrak{k}}] = -\frac{h}{2\omega}\,\delta(\mathfrak{k}-\mathfrak{k}')$$

gilt [vgl. (8)]. Beide V.-R. sind nicht ad hoc eingeführt, sondern ganz natürlich aus dem allgemeinen quantenmechanischen Formalismus entstanden. Dies genügt, wie wir sehen werden, um das richtige Vorzeichen der Gravitationswirkungen zu erhalten. Damit ist der fundamentale Unterschied zwischen Coulomb- und Newtonkräften quantenmechanisch erklärt.

Fock und Podolsky[1] haben, die Idee Diracs verfolgend, eine Herleitung des Coulombschen Gesetzes gegeben. Unsere Rechnung läuft der Fock-Podolskyschen genau parallell. Wir gehen von den Gleichungen

$$\left(\frac{1}{2m_1}\mathfrak{p}_1^2 + \frac{m_1}{2}h_{00}(\mathfrak{r}_1)\right)\psi + \frac{h}{i}\frac{\partial\psi}{\partial t_1} = 0,$$

$$\left(\frac{1}{2m_2}\mathfrak{p}_2^2 + \frac{m_2}{2}h_{00}(\mathfrak{r}_2)\right)\psi + \frac{h}{i}\frac{\partial\psi}{\partial t_2} = 0$$

aus. Bei $t_1 = t_2 = t$ haben wir

$$\left(\frac{1}{2m_1}\mathfrak{p}_1^2 + \frac{1}{2m_2}\mathfrak{p}_2^2 + \frac{h}{i}\frac{\partial}{\partial t}\right)\psi = -\left(\frac{m_1}{2}h_{00}(\mathfrak{r}_1) + \frac{m_2}{2}h_{00}(\mathfrak{r}_2)\right)\psi.$$

Die Entwicklung der Lösung nach den Potenzen von m_1 und m_2 wird gesucht.

Infolge des oben besprochenen Unterschieds in den V.-R. erhalten wir, statt der Formeln (39) und (40) von Fock und

[1] V. Fock und B. Podolsky, Sow. Phys. 1, 801, 1932 (Part II).

Podolsky, die folgenden Formeln:

$$-\left(\frac{1}{2m_1}\mathfrak{p}_1^{\,2}+\frac{1}{2m_2}\mathfrak{p}_2^{\,2}+\frac{h}{i}\frac{\partial}{\partial t}\right)\varphi_2\sim$$

$$\sim\frac{m_1}{2\,(2\pi)^{3/2}}\int h_{00,\mathfrak{k}}\,\varphi_1(\mathfrak{p}_1-h\mathfrak{k},\,\mathfrak{p}_2)\,e^{-i\omega t}\,d\mathfrak{k}+$$

$$+\frac{m_2}{2\,(2\pi)^{3/2}}\int h_{00,\mathfrak{k}}\,\varphi_1(\mathfrak{p}_1,\,\mathfrak{p}_2-h\mathfrak{k})\,e^{-i\omega t}\,d\mathfrak{k};$$

$$-\left(\frac{1}{2m_1}\mathfrak{p}_1^{\,2}+\frac{1}{2m_2}\mathfrak{p}_2^{\,2}+\frac{h}{i}\frac{\partial}{\partial t}\right)\varphi_1\sim$$

$$\sim\frac{m_1}{2\,(2\pi)^{3/2}}\int \overset{+}{h}_{00,\mathfrak{k}}\,\varphi_0(\mathfrak{p}_1+\mathfrak{k}h,\,\mathfrak{p}_2)\,e^{i\omega t}\,d\mathfrak{k}+$$

$$+\frac{m_2}{2\,(2\pi)^{3/2}}\int \overset{+}{h}_{00,\mathfrak{k}}\,\varphi_0(\mathfrak{p}_1,\,\mathfrak{p}_2+h\mathfrak{k})\,e^{i\omega t}\,d\mathfrak{k}.$$

Das liefert[1]

$$\varphi_1(\mathfrak{p}_1,\mathfrak{p}_2)=$$

$$=\frac{m_1}{2\,(2\pi)^{3/2}h^3}\,\overset{+}{h}_{00,\frac{\mathfrak{p}_1^0-\mathfrak{p}_1}{h}}\,\frac{\delta(\mathfrak{p}_2-\mathfrak{p}_2^{\,0})\,\delta_{j0}}{W-W_0+|\mathfrak{p}_1^{\,0}-\mathfrak{p}_1|}\,e^{i\frac{|\mathfrak{p}_1^0-\mathfrak{p}_1|-W_0}{h}t}-$$

$$-\frac{m_2}{2\,(2\pi)^{3/2}h^3}\,\overset{+}{h}_{00,\frac{\mathfrak{p}_2^0-\mathfrak{p}_2}{h}}\,\frac{\delta(\mathfrak{p}_1-\mathfrak{p}_1^{\,0})\,\delta_{j0}}{W-W_0+|\mathfrak{p}_2^{\,0}-\mathfrak{p}_2|}\,e^{i\frac{|\mathfrak{p}_2^0-\mathfrak{p}_2|-W_0}{h}t}.$$

Setzen wir

$$\overset{+}{h}_{00,\mathfrak{k}}\,h_{00,\mathfrak{k}}\sim 0\quad\text{und}\quad h_{00,\mathfrak{k}}\,\overset{+}{h}_{00,\mathfrak{k}}\sim\frac{h}{2\omega}\,\delta(\mathfrak{k}-\mathfrak{k}').$$

Nach der Streichung der unendlichen Selbstwirkungsglieder erhalten wir endlich

$$\left(\frac{1}{2m_1}\mathfrak{p}_1^{\,2}+\frac{1}{2m_2}\mathfrak{p}_2^{\,2}+\frac{h}{i}\frac{\partial}{\partial t}\right)\sim$$

$$\sim\frac{m_1 m_2}{4\,(2\pi)^{3/2}h}\,\frac{\delta(\mathfrak{p}_1-\mathfrak{p}_1^{\,0}+\mathfrak{p}_2-\mathfrak{p}_2^{\,0})\,\delta_{j0}}{|\mathfrak{p}_1-\mathfrak{p}_1^{\,0}|^2}\,e^{-\frac{i}{h}W_0 t}.$$

Das Vorzeichen der rechten Seite ist ein anderes als in der Fock-Podolskyschen Formel (42). Wenn wir zum Konfi-

[1] Wegen der Bezeichnungen siehe V. Fock und B. Podolsky, loc. cit.

gurationsraum zurückgehen, erhalten wir demgemäss eine Schrödingergleichung mit der potentiellen Energie

$$-\frac{m_1 m_2}{16\pi\,|\mathfrak{r}_1 - \mathfrak{r}_2|},$$

und wir haben also das Newtonsche Gravitationsgesetz als eine notwendige Folgerung der Quantentheorie der Gravitation wiedergefunden.

Physikalisch - technisches Institut
und Physikalisches Institut der Universität.
Leningrad, August 1935.

Chapter 21
Jacques Solomon (1938): Gravitation et Quanta

Jacques Solomon (1938). Gravitation et Quanta. *Journal de Physique et Le Radium*, 9: 479–485.

GRAVITATION ET QUANTA

Par Jacques SOLOMON.

Sommaire. — On étudie tout d'abord s'il est possible de parler sans contradiction de loi de Newton entre les particules élémentaires. Pour cela on discute de façon détaillée l'« expérience idéale » qui doit permettre de vérifier l'applicabilité de la loi de Newton dans le domaine considéré. On étudie ensuite les problèmes que pose la quantification du champ de gravitation et on discute enfin quelques conséquences de la théorie cosmologique présentée récemment par Dirac.

Introduction. — Pendant longtemps, gravitation et quanta ont paru deux domaines très lointains : le premier s'étendant aux phénomènes à grande échelle, le second au microscopique. La faiblesse des forces de gravitation vis-à-vis des forces électromagnétiques semblait d'ailleurs justifier cette séparation. Au reste, il apparaissait possible sans grandes difficultés (1) de généraliser le formalisme de la mécanique quantique dans un espace riemannien quelconque. Comme il n'en résultait aucune conséquence importante, toujours en raison de la petitesse des forces de gravitation, il ne semblait pas que ce sujet fût susceptible de progrès importants.

Toutefois, depuis quelque temps, la question des rapports entre gravitation et quanta a repris de l'intérêt pour deux raisons essentiellement. Tout d'abord on a cherché, suivant une suggestion de Pauli, à mettre en relation les neutrinos, dont l'intervention est nécessaire pour assurer dans les phénomènes nucléaires la conservation de l'énergie et de la quantité de mouvement, avec les ondes gravitationnelles. D'autre part, les recherches d'Eddington et de Milne, quelque discutées qu'elles soient, ont amené à examiner à nouveau la théorie quantique de la gravitation. Le but du présent travail n'est autre que de contribuer à éclaircir cette difficile question.

1. L'examen du problème des relations entre théorie de la gravitation et théorie des quanta a conduit récemment certains auteurs (2) à mettre en question la possibilité de parler de façon conséquente de forces gravifiques entre particules élémentaires. Le raisonnement est le suivant : la variation d'impulsion de l'une des particules pendant le temps de mesure Δt, dûe à la force gravifique doit être supérieure à l'imprécision « naturelle » sur l'impulsion Δp :

$$\frac{GM^2}{r^2}\,\Delta t > \Delta p > \frac{\hbar}{\Delta r} > \frac{\hbar}{r} \tag{1}$$

(où G est la constante de la gravitation de Newton, M la masse des deux particules, r leur distance, Δr l'imprécision sur celle-ci, $\hbar$ la constante de Planck divisée par 2π).

D'autre part, le déplacement des particules pendant le temps que dure la mesure doit être faible vis-à-vis de leur distance :

$$\frac{1}{2}\,\frac{GM^2}{r^2}\,(\Delta t)^2 < r \tag{2}$$

d'où, en combinant (1) et (2)

$$r > \frac{\hbar^2}{GM^3} \sim 10^{26}\ \text{cm}, \tag{3}$$

soit une distance de l'ordre du rayon de l'univers. La mesure serait donc impossible.

Si l'on applique de telles considérations aux électrons et à la loi de Coulomb, on voit qu'il serait impossible de parler de l'application de cette loi à des distances inférieures à $\frac{\hbar^2}{me^2}$, c'est-à-dire justement dans le domaine atomique. C'est ainsi par exemple, que toute la théorie de l'atome d'hélium, dont on connaît l'extrême précision, en particulier grâce aux efforts de E. Hylleraas (1) est basée sur l'hypothèse qu'entre deux électrons existe une interaction de Coulomb $\frac{e^2}{r}$. Or ces deux électrons sont en moyenne à une distance inférieure à $\frac{\hbar^2}{2me^2}$.

2. Pour résoudre cette difficulté, il importe de distinguer deux ordres de problèmes que l'on peut se poser à propos de mesure du champ :

a) Mesure du champ en un point déterminé de l'espace-temps (ou de façon plus précise, valeur moyenne du champ pour un petit élément de volume de l'espace-temps).

b) Mesure de la force s'exerçant sur une particule donnée. Ces deux problèmes ne font qu'un dans la théorie classique. Les profondes recherches de Bohr et Rosenfeld (2) nous ont montré, tout au moins dans le cas du champ électromagnétique que pour

(1) E. Schrodinger, *Berl. Ber*, 193z ; W. Pauli et J. Solomon, *J. Physique*, 1932, **3**, 582 ; V. Fock, *J. Physique*, 1929, **10**, 392 ; *cf.* cependant la remarque de L. Brillouin dans son livre sur les Tenseurs en mécanique et en élasticité, p. 201.

(2) W. Heitler, L. Nordheim, E. Teller, cités par Gamow, *Phys. Z.*, 1937, **38**, p. 814.

(1) Voir l'article de H.-A. Bethe, *Handbuch der Physik*, XXIV/I, p. 324 et suiv.

(2) N. Bohr et L. Rosenfeld, *Dansk Vidensk. Selk*, 1933, **12**, 8.

obtenir la précision optima dans la mesure, il faut justement faire usage de corps d'épreuve composés d'un grand nombre de particules élémentaires. L'utilisation d'un électron par exemple pour explorer les propriétés électromagnétiques d'une portion de l'espace n'est pas compatible avec la condition de précision maximum (¹).

Le problème se pose d'une façon toute différente lorsqu'il s'agit de vérifier qu'une particule élémentaire est soumise à une loi de forces donnée. On peut, en effet, alors suivre la particule dans son mouvement, et, comme nous allons le montrer, ceci permettra la vérification en question.

3. Soit t_0 l'instant auquel nous commençons l'examen de l'évolution du système constitué par les deux particules et soit $r(t)$ leur distance. Nous formerons donc à l'instant t_0 deux paquets d'ondes de dimensions Δr_1 et Δr_2 et le théorème d'Ehrenfest nous apprend que les lois de la dynamique classique seront applicables aux centres de gravité des deux paquets d'ondes pourvu que le champ ne varie pas trop sur l'étendue du paquet d'onde, soit en désignant de façon générale la force par F

$$\Delta r_1 \,\mathrm{grad}_1 F << F,$$

ce qui pour le cas de la loi de Newton (comme pour la loi de Coulomb), nous donne simplement la condition

$$\Delta r_1 << r. \tag{4}$$

Autrement dit, à moins que les deux trajectoires ne se rapprochent au point de voir se chevaucher les deux paquets d'onde, la mécanique classique est toujours applicable, on pourra suivre la trajectoire de chacune des deux particules, l'identifier au point de vue géométrique et en déduire de manière rationnelle l'expression exacte de la loi d'interaction qui est à son origine.

4. Examinons maintenant les restrictions que nous impose la relation (4). S'il s'agit de deux particules libres, en l'absence de toute autre interaction que leur interaction mutuelle, et si le système de référence est à peu près au repos par rapport au centre de gravité du système, on sait que la limite inférieure de Δr_1 sera approximativement $\frac{h}{Mc}$. Par conséquent, tant que les deux particules sont à une distance supérieure à

$$\frac{h}{Mc}\left(\frac{h}{Mc} = 0{,}18.10^{-13}\ \mathrm{cm}\right)$$

l'application de la loi de Newton peut être garantie. Il est par ailleurs clair qu'à ces distances même s'il s'agit de particules neutres (neutrons), les interactions autres que l'interaction de Newton sont considérablement plus importantes que celle-ci et conditionnent l'évolution du système.

(¹) Cf. J. Solomon, *J. Phys.*, 1933, **4**, 368-387.

Si les deux particules sont soumises à un autre champ (c'est par exemple le cas déjà cité des deux électrons de l'atome d'hélium), il pourrait sembler que Δr_1, prenant alors des valeurs beaucoup plus élevées, la loi de Newton (ou de Coulomb) ait son champ de validité sensiblement réduit. Par exemple pour l'atome d'hélium, Δr_1 est de l'ordre des dimensions atomiques et l'on retomberait sur les paradoxes indiqués précédemment. Mais il suffit dans ce cas d'avoir pu faire la démonstration pour des particules libres pour être certains que l'application à des particules liées de ce qui est vérifié pour des particules libres ne peut en aucun cas amener à des contradictions avec les résultats expérimentaux. On remarquera d'ailleurs que dans la théorie classique également on rencontrerait ce problème par exemple dans l'étude des électrons dans les métaux et dans bien d'autres questions.

5. Revenons maintenant à l'expérience décrite au paragraphe 3 et voyons de façon plus précise comment on peut éviter les paradoxes énoncés au début de ce travail. La condition (1) s'écrit

$$\Delta t > \frac{hr}{GM^2} \tag{5}$$

et nous impose une restriction lorsque nous nous proposons de mesurer le champ à un instant t donné : la mesure doit être d'autant plus longue que les particules considérées sont plus distantes et par suite que leur interaction est plus faible. Mais ceci ne contredit nullement notre procédé de mesure qui consiste à suivre la particule le long de sa trajectoire.

En accord avec les considérations précédentes, la condition (2) prend elle aussi un tout autre sens. Nous suivons les deux particules le long de leurs trajectoires, c'est-à-dire la grandeur de leurs déplacements pendant le temps de mesure causés par leur interaction mutuelle. De façon plus précise la relation (1) (ou (5) se rapporte à l'interaction incontrôlable qui s'introduit dans une mesure. Par contre, la relation (2) n'est nullement liée à une interaction incontrôlable ; tout au contraire, elle fait partie intégrante du processus de mesure, elle marque, comme nous venons de le dire, le déplacement le long de la trajectoire, c'est-à-dire justement ce qui fait l'objet de la mesure.

Tout autre serait évidemment la situation si nous imposions de connaître directement la force qu'exerce une particule sur l'autre en un point et à un instant donnés. Mais nous avons vu au paragraphe 2 qu'il s'agirait là d'un tout autre problème. La condition (2) par suite ne nous impose aucune restriction : elle nous indique seulement que la longueur de l'arc parcouru par la particule pendant la mesure est supérieure à

$$\frac{1}{2}\,\frac{GM^2}{r^2}\,(\Delta t)^2$$

et c'est seulement si on y joint la condition (5) que le déplacement sera supérieur à $\frac{h^2}{2GM^3}$.

6. Nous commençons à suivre nos deux particules à l'instant $t = t_0$. A ce moment nous mesurons leur quantité de mouvement $p_1(t_0)$ et $p_2(t_0)$. Soit τ la durée de cette mesure, v et v' les vitesses avant et après la mesure.

On sait [1] que l'incertitude sur p est au moins

$$\delta p = \frac{h}{(v - v')\tau}.$$

La théorie de la relativité nous montre que $v - v'$ ne saurait dépasser la vitesse de la lumière c, de sorte qu'on obtient la précision maximum sur p en posant

$$\delta p = \frac{h}{c\tau}. \tag{6}$$

A cette incertitude sur la quantité de mouvement est d'ailleurs liée une incertitude

$$\delta x = c\tau \tag{7}$$

sur la position de la particule.

D'autre part le mouvement de chaque particule est défini par

$$\frac{\mathrm{d}p_x}{\mathrm{d}t} = \frac{GM^2}{r^2}\cdot\frac{x}{r}, \tag{8}$$

d'où l'on tire par intégration le long de la trajectoire

$$p_x(t_0 + T) - p_x(t_0) = GM^2\int_{t_0}^{t_0+T}\frac{\mathrm{d}t}{r^2}\frac{x}{r}. \tag{9}$$

Notre problème consistant en la vérification de cette relation, nous mesurerons les quantités de mouvement initiale et finale, ainsi que l'intégrale du second membre. D'après (6), l'erreur commise sur le premier membre est de l'ordre de $\frac{h}{c\tau}$, alors que l'ordre de grandeur du second membre est

$$\frac{GM^2}{r^2}\,T.$$

Quant à l'erreur qu'il est possible de commettre sur l'évaluation du second membre, on peut l'évaluer d'après (7) à

$$2\,GM^2c\tau\int_{t_0}^{t_0+T}\frac{\mathrm{d}t}{r^3}\cdot\frac{x}{r} \sim \frac{2\,GM^2c\tau\,T}{r^3}.$$

L'incertitude relative sera dans le premier cas

$$\frac{hr^2}{cGM^2\tau T} \tag{10}$$

et dans le second cas

$$\frac{c\tau}{r}. \tag{11}$$

De (10) on tire la condition

$$T >> \frac{hr}{GM^2}\cdot\frac{r}{c\tau}$$

et à fortiori, d'après (11)

$$T >> \frac{r}{\frac{GM^2}{h}} \tag{12}$$

ce qui nous indique que le déplacement pendant la mesure est certainement supérieur à $\frac{h^2}{GM^3}$. Nous retombons donc sur les paradoxes que nous avions voulu éviter.

7. En fait, nous allons voir que c'est l'application brutale de formules telles que (6) qui nous a conduit à des paradoxes. L'échange incontrôlable d'énergie entre système mesuré et appareil de mesure qui caractérise l'observation dans la mécanique quantique n'a pas en effet pour conséquence comme on le dit souvent, que lorsque la position est déterminée, la quantité de mouvement est indéterminée et vice-versa. Ce qui est indéterminé, dans les limites précisées par la mécanique quantique, c'est la différence entre la quantité de mouvement (par exemple) avant la mesure et après la mesure. C'est ainsi qu'une détermination précise de la position d'une particule est parfaitement compatible avec la connaissance précise de sa quantité de mouvement à l'instant où commence la mesure. C'est la quantité de mouvement après la mesure qui est indéterminée.

La distinction bien souvent n'est pas faite parce qu'on s'intéresse à la prédiction des états ultérieurs du système, prédiction pour laquelle naturellement la connaissance de la quantité de mouvement (ou de la position) après la mesure importe seule. Mais ici, cette distinction va se révéler essentielle. Il est clair qu'elle résulte de l'applicabilité du principe de conservation de l'énergie et de la quantité de mouvement au système : particules observées + appareil de mesure.

Par exemple, nous pourrons en principe mesurer notre quantité de mouvement puis, après un intervalle de temps arbitrairement petit, mesurer la position : l'erreur commise est au plus $c\tau_1$, et peut-être rendue aussi petite que l'on veut. On aura ainsi mesuré la quantité de mouvement avant la mesure de la position qui pourra être aussi précise que l'on veut et entraînera une incertitude en conséquence sur la quantité de mouvement *après* la mesure.

Si maintenant nous considérons la quantité de mouvement à la fin de notre expérience sur la gravitation $p(t_0 + T)$, il est clair que seule nous intéresse la valeur de p *avant* la détermination de celle-ci, c'est-à-dire au début de l'intervalle de temps τ. C'est ce que l'on peut réaliser par exemple, en disposant au point considéré un écran macroscopique

[1] *Cf.* p. ex. W. Pauli, *Handbuch der Physik*, XXIV/I, p. 93.

de masse $\mathfrak{M}$. On pourra observer le recul de celui-ci au moment du choc avec toute la précision désirable et par suite $p\ (t_0 + T)$ sans aucune incertitude.

Au contraire, pour la détermination de $p\ (t_0)$, c'est de sa valeur *après* la mesure que nous avons besoin et il est clair qu'il nous est impossible d'utiliser le même raisonnement.

8. Imaginons un instant qu'il s'agisse de mettre en évidence l'interaction de Coulomb entre deux particules. On utilisera à cet effet le dispositif de la figure 1.

Les deux particules pénètrent par des orifices de dimensions linéaires Δx dans l'appareil de mesure

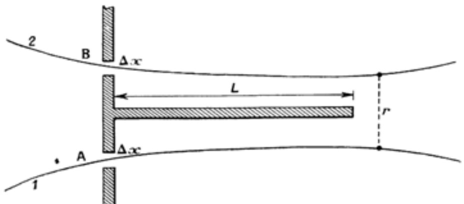

Fig. 1.

et sont séparés l'un de l'autre par une cloison de longueur L. Nous savons donc que la particule 1 à l'instant t_0 était en A, sa position étant connue à Δx près d'où une indétermination sur la différence des quantités de mouvement avant et après le passage par A qui est de l'ordre de $\frac{\hbar}{\Delta x}$. Nous mesurons *ensuite* pendant un temps suffisamment long la quantité de mouvement de 1 (par exemple par effet Doppler). Naturellement, pendant cette mesure, nous ignorons complètement la position de la particule, mais si la mesure dure le temps τ_1, nous connaîtrons la quantité de mouvement à $\frac{\hbar}{c\tau_1}$ près. D'après le théorème d'Ehrenfest, connaissant la position de 1 à l'instant t à Δx près, connaissant d'autre par la quantité de mouvement, nous serons en mesure de calculer à tout instant la position de la particule 1 tant qu'elle n'est soumise à aucun champ extérieur. Soit v la vitesse de la particule avant son interaction avec l'autre. τ_1 sera au plus égal à $\frac{L}{v}$. L'incertitude relative calculée au paragraphe 6 ne fait donc plus intervenir le temps très court $\tau \sim \frac{\Delta x}{c}$ mais le temps $\tau_1 = \frac{L}{v}$ qui peut-être rendu beaucoup plus long si l'on prend v suffisamment petit (quoique naturellement subsiste la relation

$$Mv > \frac{\hbar}{\Delta x})$$

ou L suffisamment grand.

9. En passant de l'électromagnétisme à la gravitation, on notera que le raisonnement précédent est basé sur la considération d'un appareil qui permet d'exclure de manière complète l'interaction entre les deux particules dans une région déterminée de l'espace. Il semblerait donc nécessaire de faire usage dans notre problème de la notion de paroi imperméable à la gravitation. Supposant admise l'existence de telles parois, nous allons vérifier que nous avons obtenu la solution du problème étudié.

L'incertitude relative calculée au paragraphe 6 n'est plus donnée par

$$\frac{\hbar r^2}{cGM^2\tau T} \tag{10 bis}$$

mais par

$$\frac{v}{c}\cdot\frac{r}{L}\cdot\frac{r}{\frac{GM^2}{\hbar}T} \tag{13}$$

d'où la condition

$$T > \frac{v}{c}\cdot\frac{r}{L}\cdot\frac{r}{\frac{GM^2}{\hbar}}$$

qui nous montre que le déplacement pendant la mesure proprement dite (temps T) sera supérieur à

$$\frac{1}{2}\,\frac{v^2}{c^2}\cdot\frac{\hbar^2}{GM^2}\cdot\frac{r^2}{L^2}$$

et la condition (3) est par suite remplacée par

$$r < L.\frac{L}{\frac{\hbar^2}{GM^2}}.\frac{c^2}{v^2} < L.\frac{L}{\frac{\hbar^2}{GM^2}}\left\{\frac{\Delta x}{\frac{\hbar}{Mc}}\right\}^2. \tag{14}$$

Tout d'abord cette condition nous montre que r n'est pas limité supérieurement mais inférieurement. On notera ensuite qu'elle nous fournit des ordres de grandeur tout à fait raisonnables : il est nécessaire en effet comme nous l'avons remarqué au début du paragraphe 4 que r soit supérieur à $\frac{\hbar}{Mc}$, d'où une condition pour L :

$$L^2 >> \left\{\frac{\frac{\hbar}{Mc}}{\Delta x}\right\}^2\cdot\frac{\hbar}{Mc}\cdot\frac{\hbar^2}{GM^2}, \tag{15}$$

soit numériquement

$$L >> \left\{\frac{\frac{\hbar}{Mc}}{\Delta x}\right\}.10^6\ \text{cm}.$$

En prenant un écran matériel, on aura

$$\Delta x \sim \frac{h^2}{me^2} \sim 10^{-8} \text{ cm},$$

d'où avec $\frac{h}{Mc} = 2.10^{-14}$ cm, $L >> 1$ cm,

ce qui est très raisonnable.

On pourra par exemple partir des données suivantes :

$$\Delta x = 10^{-8} \text{ cm},$$
$$L = 2.10^4 \text{ cm},$$
$$r = 10^{-6} \text{ cm},$$
$$\frac{L}{v} = \frac{1}{3} \text{ sec},$$
$$T > 10^{-48} \text{ sec}.$$

Pour que le déplacement soit de l'ordre de $r = 10^{-6}$ cm, il faudrait que T soit de l'ordre de 10^{10} siècles (3.10^{18} sec). Pendant ce temps, le parcours dû à l'impulsion initiale est de l'ordre de 10^{23} cm, donc très inférieur au rayon de l'univers.

On pourrait peut-être encore objecter aux considérations qui précèdent qu'un dispositif comme celui de la figure 1 rend difficile la détermination de l'instant où commence l'interaction newtonnienne entre les deux particules. Mais, vu la longue durée de l'observation, l'erreur possible sur l'instant initial est négligeable. Et l'on observera d'autre part, que du fait que nous nous en sommes tenus à l'approximation newtonienne de la loi de la gravitation, il est tout à fait conséquent de négliger les erreurs dûes au fait que l'interaction entre les deux particules ne se propage pas de manière instantanée, mais avec la vitesse de la lumière.

X. Il nous reste à examiner l'emploi qui a été fait d'écrans opaques à la gravitation. On sait qu'il n'existe pas de tels écrans dans le sens où l'on parle d'écrans opaques aux actions électromagnétiques. Toutefois si l'on s'imagine un écran matériel d'une certaine épaisseur, cet écran va entrer en interaction avec les deux particules et perturber leurs trajectoires. Comme la masse de l'écran est infiniment grande vis-à-vis de chacun des deux particules, le fait que l'une des particules est présente ou absente dans l'une des moitiés de l'appareil n'influe pratiquement en rien la trajectoire de l'autre particule, de sorte que tout se passe comme si nous avions réellement un écran opaque à la gravitation.

Naturellement, le fait que l'écran modifie la trajectoire de deux particules complique quelque peu le schéma de notre expérience, mais n'en altère nullement les conclusions. Il suffira en effet, de faire deux fois l'expérience : une fois avec une particule seulement et la seconde fois avec les deux particules pour pouvoir par différence mettre en évidence l'interaction newtonienne entre les deux particules.

XI. Les temps et les longueurs qui sont nécessaires pour mettre à exécution l'expérience précédente peuvent sembler de nature à jeter un doute sur la possibilité de tels processus de mesure. Il faut toutefois rappeler que les interactions de gravitation sont extraordinairement petites : à distance égale l'attraction newtonienne de deux protons est 10^{36} fois plus petite que leur répulsion coulombienne, il faudra donc pour la mettre en évidence des temps extrêmement longs. D'autre part, aucune limitation de longueur ou de durée n'est inhérente à la théorie elle-même. La notion de rayon de l'univers elle-même doit sans doute être mise en relation avec la masse totale de l'univers et il est peu vraisemblable qu'elle puisse avoir quelque chose à faire avec le problème de deux particules. L'essentiel dans la construction d'une telle expérience « idéale » est qu'elle soit possible sans contradiction avec les résultats théoriques connus et c'est ce qui semble bien résulter des développements précédents.

XII. Nous nous sommes placés jusqu'ici dans le cadre de la théorie de Newton. On peut examiner ce que devient la notion de champ de gravitation dans le domaine de la théorie générale de la relativité.

Le problème de la quantification du champ de gravitation, qui a déjà donné lieu à un certain nombre de recherches donne lieu à d'importantes difficultés. Tout d'abord les équations de gravitation de la théorie générale de la relativité ne sont pas linéaires comme les équations de Maxwell. Dans le cas où l'on peut se borner à considérer les écarts entre le champ de gravitation et un champ de Minkowski comme faibles, le problème peut être résolu par les méthodes analogues à celles de la quantification du champ électromagnétique (1).

Dans ce cas, qu'on peut appeler « cas des champs de gravitation faibles », Bronstein (2) a étudié les possibilités de mesure des grandeurs de champ, c'est-à-dire dans le cas présent des symboles à trois indices de Christoffel. On utilise à cet effet un corps d'épreuve de volume V, de densité ρ. Si T est la durée de la mesure, l'erreur commise sur la moyenne d'un symbole tel que [00,1] par exemple sur le volume V et le temps T est donnée par

$$\Delta\,[00,1] \geqslant \frac{h^{2/3}\,G^{1/3}}{c^{1/3}\rho^{1/3}\,V^{2/3}\,T}. \tag{16}$$

Ceci montre que dans le domaine considéré, où les champs de gravitation sont faibles, où le principe de superposition est valable, il est toujours possible de se donner un corps d'épreuve de densité suffisamment élevée pour que l'erreur commise sur le symbole [00,1] lors de sa mesure soit inférieure à toute limite donnée. Dans ce domaine par conséquent,

(1) L. Rosenfeld, *Z. Physik*, 1930, **65**, 589-600 ; J. Solomon, *Z. Physik*, 1931, **71**, 162.

(2) M. Bronstein, *Sow. Phys.*, 1936, **9**, 140-157.

il y a accord entier entre les principes de la théorie de la gravitation et de la théorie des quanta et il est possible de construire une théorie quantique conséquente de la gravitation.

En particulier, comme l'a montré Bronstein, de même que d'après Dirac, l'interaction de Coulomb entre deux charges électrisées provient des possibilités d'absorption et d'émission de quanta d'énergie électromagnétique, l'interaction de Newton entre deux particules douées de masse provient des possibilités d'émission et d'absorption par ces particules de quanta d'énergie gravifique. On retrouve en effet ainsi la loi de gravitation de Newton comme conséquence nécessaire de la théorie quantique de la gravitation, en bon accord avec nos conclusions sur la possibilité effective qu'il y a de considérer l'interaction de Newton entre deux particules élémentaires.

XIII. Si maintenant les champs de gravitation ne sont pas faibles, si l'écart avec l'euclidicité est très important, le problème se pose différemment. Tout d'abord on pourrait remarquer que le rayon gravitationnel du corps servant à la mesure doit, pour que la mesure ait un sens, être inférieur à ses dimensions linéaires :

$$\frac{G\rho V}{c^2} < V^{1/3},$$

d'où une limite supérieure pour ρ et, si l on en tient compte dans (16) une limitation sur la mesure des crochets [00,1].

Mais on peut se borner à remarquer que dans les cas où le champ de gravitation n'est pas faible, la méthode même de quantification, basée sur le principe de superposition, fait défaut, de telle sorte qu'il n'est plus possible d appliquer une relation telle que (16) dans un sens dépourvu d ambiguïté.

Il est d'ailleurs intéressant à ce point de vue de rappeler que Rosen (¹) a montré que les équations rigoureuses de la relativité générale n'admettent pas de solutions représentant des ondes planes polarisées d'amplitude finie qui soient en même temps dépourvues de singularités. Si le ds^2 a la forme

$$ds^2 = -A\,dx^2 - B\,dy^2 - C\,dz^2 + A\,dt^2$$

et qu'on pose

$$\sigma = \sqrt{BC},$$

si une onde de courte durée (électromagnétique ou gravifique) passe en un point donné de l'espace, on montre qu'à ce point σ commence à décroître et continue à décroître après que l'onde a passé jusqu'à ce qu'ait lieu la « catastrophe », $\sigma = 0$. Au contraire (²) il est possible de considérer des ondes cylindriques dépourvues de singularités.

(¹) N. Rosen, *Sow. Phys.*, 1937, **12**, 366-373.
(²) A. Einstein et N. Rosen, *J. Frankl. inst*, 1937, **43**, 223.

De telles considérations sont de nature à mettre sérieusement en doute la possibilité de concilier le formalisme *présent* de la quantification des champs avec la théorie *non linéaire* de la gravitation.

XIV. Dans un travail récent, Dirac (¹) a cherché à donner de nouvelles bases à la cosmologie. Pour cela, il part de la remarque que « l'âge » de l'univers, dans la théorie de l'univers, est de l'ordre de 2.10^9 années; évalué en unités électroniques $\frac{e^2}{mc^3}$, il est donc égal à 7.10^{38}, soit du même ordre que le rapport de la force électrostatique à la force gravitationnelle entre électron et proton qui est $\gamma = 2{,}3.10^{39}$. Ceci posé, il admet que ce n'est pas là une coïncidence fortuite, mais que cette égalité approximative doit être conservée dans le cours de l'évolution de l'univers. On doit même généraliser ce fait et poser en principe que deux nombres sans dimensions très grands intervenant dans la nature sont reliés par une relation mathématique simple dans laquelle les coefficients sont de l'ordre de grandeur de l'unité. On suppose à cet effet que les nombres $\frac{hc}{e^2}$ et $\frac{M}{m}$ sont constants. De là Dirac tire un certain nombre de conséquences sur lesquelles je ne veux pas insister. Je me bornerai ici à deux remarques.

Tout d'abord nous avons vu que l'on peut associer à la théorie de la gravitation la longueur

$$\frac{h^2}{GM^3} = 10^{26}\ \text{cm} \tag{17}$$

de la même façon exactement qu'on associe à la théorie coulombienne de l'atome la longueur

$$\frac{h^2}{me^2} = 10^{-8}\ \text{cm} \tag{18}$$

et que (17) est très voisin du rayon de l'univers d'après Einstein (²) ($R = 6{,}9.10^{27}$ cm). Cette coïncidence a frappé un certain nombre d'auteurs (³) qui ont voulu en tirer les principes d'une théorie « atomique » de l'Univers. Il semble néanmoins plus raisonnable d'admettre que le rayon de l'univers est en relation avec le nombre total de particules (ou avec la masse totale) contenues dans l'univers. Dans ces conditions, de deux choses l'une : ou M (ou m) est en relation avec la masse totale de l'Univers ou G est en relation avec cette masse totale. Dirac élimine la première éventualité (théorie de Mach) pour supposer que G varie de façon inversement proportionnelle au temps écoulé depuis l'origine de l'univers (7.10^8 années seulement au moment présent)

(¹) P. A. M. Dirac, *Proc. Roy. Soc.*, 1938, **165 A**, 199-208.
(²) *Cf.* p. ex. H. Mineur, L'Univers en expansion (n° 63 des *Actualités scientifiques et industrielles*).
(³) E. Schrodinger, *Nature*, **141**, 410 ; H. Ertel, *Naturwiss*, 1936, **32**, 36, 70 ; Ira M. Freeman, *Naturwiss*, 1936, **23**, 557.

Ainsi alors que la longueur (17) est proportionnelle à l'âge de l'univers et s'annule même en toute rigueur à l'origine des temps, les dimensions de l'atome sont indépendants de « l'époque ». Ceci a pour conséquence qu'à l'origine des temps, la matière est concentrée dans un espace très réduit d'où elle diffuse dans la suite : image où se rejoignent les théories, si opposées par ailleurs, semble-t-il, d'Eddington et de Milne [1].

D'autre part, on sait que la théorie de Fermi introduit dans le mécanisme de la désintégration β une constante nouvelle

$$G_1 = 3.10^{-13}$$

qui semble jouer un rôle essentiel dans la théorie des noyaux atomiques [2]. D'après le principe de Dirac, τ désignant l'époque, on doit avoir

$$G_1^3 \text{ prop. à } \frac{1}{\tau}$$

puisque $G_1^{-3} = 4.10^{40}$. On remarquera d'ailleurs que

$$G_1^{-3} = 18\,\gamma = 18\,\frac{e^2}{GMm}$$

(on sait les tentatives qui ont été faites pour relier les phénomènes de désintégration β aux phénomènes gravifiques et qui out pour base en somme la relation qui précède).

Dans ces conditions on sait que l'évaluation de l'époque actuelle à 7.10^8 années est beaucoup trop faible eu égard aux données radioactives sur l'âge de la terre. On peut chercher à éliminer la difficulté en supposant que la loi de désintégration elle-même a varié, autrement dit que ce que nous appelons la « constante de désintégration » est fonction du temps. Dans le cas de la désintégration β, la constante de désintégration est proportionnelle à G_1^2, donc à $\tau^{-2/3}$.

Au contraire, dans le cas de la désintégration α aucune nouvelle constante n'intervient en dehors des constantes universelles e, h, m, M, c, de telle sorte qu'il semble difficile d'imaginer une dépendance de la constante de désintégration α de l'époque τ. Mais même si cela était, on remarquera que si l'on prend la plus petite constante de désintégration α, soit celle du thorium ($1{,}33.10^{-18}$), en appliquant la règle de Dirac, on voit qu'elle doit être proportionnelle à $\tau^{-1/2}$. Donc les constantes de désintégration α et β varient comme $\tau^{-1/2}$ et $\tau^{-2/3}$ donc de façon très différente. Dans une branche comme celle qui a lieu après le thorium C,

$$\mathrm{Th}C \begin{array}{l} \overset{\beta}{\nearrow} \mathrm{Th}C' \\ \quad 65 \text{ p.100} \\ \quad 35 \text{ p.100} \\ \underset{\alpha}{\searrow} \mathrm{Th}C'' \end{array}$$

le rapport $\frac{\lambda_\alpha}{\lambda_\beta}$ qui est actuellement $\frac{35}{65}$ aurait donc varié jusqu'ici comme $\tau^{1/6}$, c'est-à-dire serait 30 fois plus grand qu'à l'origine. Il faudrait donc non seulement supposer que les constantes de désintégration se sont considérablement modifiées (comme $\tau^{-1/2}$ et $\tau^{-2/3}$ respectivement), mais ceci entraînerait d'importantes modifications dans les chaînes radioactives, en favorisant considérablement les émetteurs β aux dépens des émetteurs α. En tout cas, si cette conception s'avérait exacte, les déterminations de l'âge de la terre basées sur l'utilisation d'éléments radioactifs différents devraient donner des résultats différents (il suffirait que dans chaque chaîne le nombre de désintégrations α et β ne soit pas le même).

XV. **Conclusion.** — Il semble résulter de cette étude tout d'abord qu'il est possible de bâtir une théorie quantique de la gravitation au moins dans le cas des champs faibles, qu'il est par exemple possible encore de parler de la loi de Newton dans le domaine quantique sans tomber dans des contradictions. Lorsque les champs sont intenses, des difficultés se présentent pour lesquelles il ne semble pas que la théorie quantique des champs soit en mesure de présenter une solution. Les tentatives cosmologiques qui ont été présentées depuis quelque temps en vue de présenter une théorie quantique de l'univers dans son ensemble sont discutées. De nombreuses difficultés s'opposent à ce que l'on puisse construire une théorie dans laquelle seule, en somme, la constante de gravitation G est mise en relation avec l'univers dans son ensemble.

Manuscrit reçu le 20 août 1938.

[1] *Cf.*, A. Lemaître, *Nature*, 1931, **127**, 706.

[2] W. Heisenberg, *Z. Physik*, 1936, **101**, 533.

Chapter 22
Paul Weiss (1938): On the Hamilton-Jacobi Theory and Quantization of a Dynamical Continuum

Paul Weiss (1938). On the Hamilton-Jacobi Theory and Quantization of a Dynamical Continuum. *Proceedings of the Royal Society of London*, A169, 102–119.

On the Hamilton-Jacobi theory and quantization of a dynamical continuum

By P. Weiss, Ph.D., *Downing College, Cambridge*

(*Communicated by R. H. Fowler, F.R.S.—Received* 1 *September* 1938)

1. Introduction

The quantization of the dynamics of point systems is closely connected with the Hamilton-Jacobi theory of the calculus of variations for *simple* integrals, the latter being a suitable mathematical formalism for describing the classical laws of point dynamics. This connexion is evident from the fact that, in order to quantize a dynamical system, one has first to know what the *pairs of canonically conjugate variables* are, and also from the fact that the quantum laws, if expressed in terms of commutation brackets, have exactly the same form as the classical laws, if expressed in terms of *Poisson brackets*.

In dealing with the quantization of the dynamics of continuous media, e.g. the electromagnetic and other fields, which has been developed along different lines, following Heisenberg and Pauli (1929), it seems tempting to try to base the method of quantization on an extended Hamilton-Jacobi theory of the calculus of variations for *multiple* integrals, the latter being the appropriate formalism for describing most relevant systems of continuum physics.

In a previous paper (Weiss 1936), referred to as I, such an attempt has been made, by extending the notion of pairs of conjugate variables. That attempt led to quantum relations on an arbitrary space-like section in space-time, provided that one *postulated* that, if the space-like section becomes especially a space section, the quantum relations should go over into those of Heisenberg and Pauli.

In the present paper the notion of Poisson brackets will be extended as well, and it will be *deduced* that the classical laws of a dynamical continuum (which satisfies a variation principle), expressed in terms of Poisson brackets, have exactly the same form as the quantum laws arrived at in I. The procedure of quantizing a dynamical continuum therewith becomes exactly the same as the procedure of quantizing point dynamics, and it will yield all physically relevant results of I.

The main part of this paper will be concerned with classical considerations, i.e. the Hamilton-Jacobi theory of continua leading from Euler's partial differential equations to Poisson brackets. The transition to the quantum theory of continua will then be a matter of a few lines. The argument, which is modelled on the Hamilton-Jacobi theory of mechanics, starts from the "complete variation" and "boundary formula", like the argument of I, and will then lead on to an extension of Poincaré's "relative integral invariant", from there to Lagrange brackets and then to Poisson brackets.

To obtain Poincaré's integral invariant and the subsequent notions, it will be necessary to use a representation of all quantities in terms of *initial data*. The use of such a representation presupposes that the state of the dynamical continuum in space-time is uniquely determined by initial data (plus boundary data in space) and that it depends continuously on these data. In other words, the initial value problem must be *correctly set*.* This is true only for theories describing a *dynamical* continuum, not, for instance, for equilibrium theories. Thus we have the satisfactory feature that the method of quantization is for purely mathematical reasons restricted to dynamical theories, as it should be on obvious physical grounds.

The initial data need not be given on a space section ($t = \text{const.}$), but may be given on any space-like section. Mathematically, this is well known from the theory of partial differential equations. The physical implications of this fact are that the present formalism will be relativistically invariant throughout. This is an advantage as compared with the work of Heisenberg and Pauli, where a space section was fixed once and for all, in consequence of which the relativistic invariance of the result was not apparent and had to be proved separately.

Another implication relates to the notion of "state" in continuum dynamics. The word "state" in the classical sense (cf. Cartan 1922, p. 4) is synonymous with "correctly set data", and it is clear from the above remarks that it will be relativistic. It is also clear, from the meaning of the word "data", that though referring only to a space-like cross-section, it describes the motion of the dynamical continuum throughout space-time. The same is true for the word "state" in the quantum sense (Dirac 1935, ch. I), meaning the maximum information of non-contradictory data relating to the dynamical system. For the fundamental difference between the classical and the quantum sense of the word state does not affect the question of its relativistic invariance. Thus by allowing space-like sections

* For a detailed discussion of the notion of "correctly set data" in the theory of partial differential equations, cf. Hadamard (1923, 1932, ch. I) and Courant-Hilbert (1937, chs. III, V).

as bearers of the data, and not restricting oneself to space sections of constant time, the two meanings of the word state, distinguished by Dirac both in classical and quantum physics, can be reconciled with each other.

Some basic considerations of I will be repeated. As in I, the continuum will be n-dimensional and will be described by ν dependent variables. The notions of "space-like" and "time-like" can still be defined mathematically for an n-dimensional space-time, through the theory of differential equations of hyperbolic type. This generality serves to accentuate the mathematical character of our procedure. The comparison with the case of mechanics, $n=1$, will be discussed at different stages of the argument. The mathematical tool applicable to the general case, $n>1$, is furnished by the theory of functionals.*

The present considerations are not immediately applicable to physics, since all physical continuum theories, like electrodynamics, have special features requiring special discussions.† We shall deal with these questions in a later communication where the present considerations will serve as a suitable basis.

2. Action integral

We consider a system with n independent variables x^i and ν dependent variables z^α, whose classical laws satisfy a variation principle. The derivatives $\frac{\partial z^\alpha}{\partial x^k}$ will be denoted by z^α_k. Let

$$\mathsf{I} = \int_D^{(n)} \mathscr{L}\{x^i, z^\alpha(x), z^\alpha_k(x)\}\, dx \quad \begin{pmatrix} i, k = 1, 2, \ldots, n \\ \alpha = 1, 2, \ldots, \nu \end{pmatrix} \tag{2·1}$$

be the action integral.‡ It is an n-fold integral extended over a domain D of the n-dimensional "x-space". dx stands for the product $dx^1 dx^2 \ldots dx^n$. The boundary of D, which forms an $(n-1)$-dimensional closed manifold,

* Cf. Volterra (1913, 1930, 1937), De Donder (1935), Prange (1915), and Juvet (1926). The present considerations resemble more closely Prange's work than that of any other author. Many of the notions used here have been defined by Prange for the case of a two-dimensional continuum. He did not, however, consider the question of correctly set data, so that his arguments referring to Poincaré's integral invariant, etc. are merely formal and his analogies with mechanics not quite correct.

Poisson brackets have first been introduced in the theory of functionals by Volterra (1913, p. 74), but without reference to the calculus of variations.

† In I, § 5, the special features of electrodynamics and their consequences have been discussed in detail.

‡ Throughout this paper Greek indices will vary from 1 to ν, Latin indices $i, j, k, \ldots$ from 1 to n, Latin indices $r, s, \ldots$ from 1 to $n-1$. We shall use the customary summation convention.

will be denoted by S and be referred to $n-1$ parameters u^r. The integrand $\mathscr{L}$ is the Lagrangian function which characterizes the particular system under consideration.

In the case $n=1$ and if x denotes the time, we have a mechanical system of ν degrees of freedom. The boundary S degenerates in this case into two distinct points and the action integral becomes

$$\mathsf{I} = \int_1^2 \mathscr{L}\{x, z^\alpha(x), z'^\alpha(x)\}\, dx. \tag{2·1a}$$

This difference between the cases $n=1$ and $n>1$ is responsible for some difficulties which arise if one tries to generalize the theory of Hamiltonian mechanics to the case of continuous media.

It will sometimes be convenient to picture the functions $z^\alpha(x)$ as an n-dimensional surface in the $(n+\nu)$-dimensional "(x,z)-space". Its $(n-1)$-dimensional closed boundary may still be referred to the parameters u^r of S; for S is the projection of that boundary into the x-space.

3. Complete variation

We define variations η^α by varying the functions $z^\alpha(x)$,

$$\bar{z}^\alpha(x) = z^\alpha(x) + \epsilon\eta^\alpha(x), \tag{3·1}$$

and variations δx^i, δz^α by

$$\bar{x}^i = x^i + \epsilon\,\delta x^i(x), \quad \bar{z}^\alpha(\bar{x}) = z^\alpha(x) + \epsilon\,\delta z^\alpha(x), \tag{3·2}$$

ϵ being an infinitesimal parameter.

These variations are connected by the relation

$$\delta z^\alpha = \eta^\alpha + z^\alpha_l\,\delta x^l. \tag{3·3}$$

The derivatives $\dfrac{\mathfrak{d}\eta^\alpha}{\mathfrak{d}x^k}$ will be denoted by η^α_k.

If we picture our quantities in the (x,z)-space, the variations δx^i, δz^α are more fundamental than the η^α, because $(\epsilon\,\delta x^i, \epsilon\,\delta z^\alpha)$ is the displacement of the point (x^i, z^α) in the (x,z)-space.

By the "*complete variation*" $\delta\mathsf{I}$ we understand the variation of the action integral I due to the variations (3·1) and (3·2). We obtain for it

$$\delta\mathsf{I} = \int_D^{(n)} \left(\frac{\partial\mathscr{L}}{\partial z^\alpha}\eta^\alpha + \frac{\partial\mathscr{L}}{\partial z^\alpha_k}\eta^\alpha_k\right) dx + \oint_S^{(n-1)} \mathscr{L} N_i\,\delta x^i\, du,$$

where the N_i are the $(n-1)$-rowed minors of the Jacobian matrix*

$$\left\| \frac{\mathfrak{d} x^i}{\mathfrak{d} u^r} \right\|;$$

they form a vector field on S, normal to S. du stands for the product $du^1 du^2 \ldots du^{n-1}$.

With the help of the formula

$$\frac{\partial \mathscr{L}}{\partial z_k^\alpha} \eta_k^\alpha = -\left(\frac{\mathfrak{d}}{\mathfrak{d} x^k} \frac{\partial \mathscr{L}}{\partial z_k^\alpha}\right) \eta^\alpha + \frac{\mathfrak{d}}{\mathfrak{d} x^k}\left(\frac{\partial \mathscr{L}}{\partial z_k^\alpha} \eta^\alpha\right)$$

we perform an integration by parts, using Green's theorem, and introduce on the boundary the variations δx^i, δz^α instead of η^α, by means of (3·3). $\delta \mathsf{I}$ then becomes

$$\delta \mathsf{I} = \int_D^{(n)} \left(\frac{\partial \mathscr{L}}{\partial z^\alpha} - \frac{\mathfrak{d} p_\alpha^l}{\mathfrak{d} x^l}\right) \eta^\alpha dx + \oint_S^{(n-1)} (X_i \delta x^i + P_\alpha \delta z^\alpha)\, du, \tag{3·4}$$

where we have introduced the notations†

$$p_\alpha^i = \frac{\partial \mathscr{L}}{\partial z_i^\alpha}, \quad U_k^i = \mathscr{L} \delta_k^i - p_\beta^i z_k^\beta, \tag{3·5}$$

$$P_\alpha = p_\alpha^l N_l, \quad X_k = U_k^l N_l. \tag{3·6}$$

The ν coefficients of η^α under the n-fold integral on the right-hand side of (3·4) are the Eulerian expressions. If they vanish, we have the Eulerian equations

$$\frac{\partial \mathscr{L}}{\partial z^\alpha} - \frac{\mathfrak{d} p_\alpha^l}{\mathfrak{d} x^l} = 0. \tag{3·7}$$

In physics these equations constitute the classical laws. Their solutions are the "extremals" of the calculus of variations.

In the case of mechanics, $n = 1$, the Eulerian equations are the equations of motion. In the case of a continuum, $n > 1$, they are one set of field equations, another set being given by the integrability conditions which the $z_k^\alpha(x)$ must satisfy in order to be representable as derivatives $\frac{\mathfrak{d} z^\alpha}{\mathfrak{d} x^k}$; these read

$$\frac{\mathfrak{d} z_k^\alpha}{\mathfrak{d} x^i} - \frac{\mathfrak{d} z_i^\alpha}{\mathfrak{d} x^k} = 0. \tag{3·8}$$

From now onwards we shall assume that (3·7) and (3·8) are satisfied.

* The quantities N_i have in I been denoted by X_i. We have changed our notation in a few details in the hope that the present notation will be more suggestive.

† The quantities X_i have in I been denoted by σ_i and the P_α by π_α.

4. Boundary formula and conjugate variables

It follows from (3·4) that extremals, instead of being characterized by (3·7), may alternatively be characterized by the equation

$$\delta \mathsf{I} = \oint_S^{(n-1)} (X_i \delta x^i + P_\alpha \delta z^\alpha)\, du. \tag{4·1}$$

We shall call this equation the "*boundary formula*".* The action integral I may be considered as a functional of the boundary—$x^i(u)$, $z^\alpha(u)$—in the (x,z)-space and the boundary formula (4·1) as its differential. X_i and P_α are therefore the functional derivatives of I with respect to x^i and z^α respectively, taken at a point (u) of the boundary. By this property X_i and P_α are defined as the conjugates to x^i and z^α so that we have the "*pairs of conjugate variables*"

$$(X_i, x^i), \quad (P_\alpha, z^\alpha). \tag{4·2}$$

In the case of mechanics the boundary formula becomes

$$\delta \mathsf{I} = \left| \mathscr{H}\, \delta x + p_\alpha \delta z^\alpha \right|_1^2, \tag{4·1a}$$

where we have put†

$$p_\alpha = \frac{\partial \mathscr{L}}{\partial z'^\alpha}, \quad \mathscr{H} = \mathscr{L} - p_\beta z'^\beta \tag{3·5a}$$

and inverted the sign of the variations at the point 1. $\mathscr{H}$ represents the negative energy of the system. It will be noticed that (3·5) and (3·6) *both* reduce to (3·5a).

I degenerates into a "two-point function" of the co-ordinates (x^i, z^α) of the two boundary points in the (x,z)-space. The pairs

$$(\mathscr{H}, x), \quad (p_\alpha, z^\alpha) \tag{4·2a}$$

are the well-known conjugate variables of Hamiltonian mechanics. This fact shows that our present considerations may be considered as a generalization of Hamiltonian mechanics.

It should be noted that in the general case, $n > 1$, the conjugates of x^i and z^α respectively depend not only on a point in the x-space, like the z^α, but also on the *direction* of a surface element passing through this point. This fact will be important in our subsequent considerations.

* This term is the translation of the French "formule aux limites". Cf. Hadamard (1910, pp. 142–51).

† Our definition of $\mathscr{H}$ differs from the usual one in sign.

5. Natural co-ordinate system and Legendre transformation

A "*natural co-ordinate system*" is defined in the x-space and is adapted to the boundary S in such a way that on S the first co-ordinate, say w, is normal to S and the other $n-1$ co-ordinates are the parameters u^r of S. We shall use the notation

$$z'^{\alpha} = \frac{\mathfrak{d} z^{\alpha}}{\mathfrak{d} w}, \quad \tilde{z}^{\alpha}_r = \frac{\mathfrak{d} z^{\alpha}}{\mathfrak{d} u^r}. \tag{5·1}$$

For the conjugates P_α, X_k, defined by (3·6), we obtain in a natural co-ordinate system

$$P_\alpha = \frac{\partial \mathscr{L}}{\partial z'^{\alpha}}; \quad X_1 = \mathscr{L} - P_\beta z'^{\beta}, \quad X_{1+r} = -P_\beta \tilde{z}^{\beta}_r. \tag{5·2}$$

Let us put
$$\mathscr{H} = \mathscr{L} - P_\beta z'^{\beta}, \quad h_r = -P_\beta \tilde{z}^{\beta}_r. \tag{5·3}$$

The boundary formula (4·1) may now be written as

$$\delta \mathsf{I} = \oint_S^{(n-1)} (\mathscr{H}\,\delta w + h_r \delta u^r + P_\alpha \delta z^\alpha)\, du. \tag{5·4}$$

By means of the *Legendre transformation* (Courant-Hilbert 1937, p. 26)

$$z'^{\alpha} \to P_\alpha, \quad \mathscr{L}(z^\alpha, \tilde{z}^{\alpha}_r, z'^{\alpha}) \to \mathscr{H}(z^\alpha, \tilde{z}^{\alpha}_r, P_\alpha) \tag{5·5}$$

we replace the derivatives z'^{α} normal to S by the P_α, and the classical laws can now be written in the "*canonical form*"

$$\frac{\partial \mathscr{H}}{\partial z^\alpha} - \frac{\mathfrak{d}}{\mathfrak{d} u^r}\frac{\partial \mathscr{H}}{\partial \tilde{z}^{\alpha}_r} - \frac{\mathfrak{d} P_\alpha}{\mathfrak{d} w} = 0, \quad \frac{\partial \mathscr{H}}{\partial P_\alpha} + \frac{\mathfrak{d} z^\alpha}{\mathfrak{d} w} = 0. \tag{5·6}$$

In the case $n=1$ the question of the natural co-ordinate system does not arise, $\mathscr{H}$ goes over into the Hamiltonian of mechanics, and the canonical equations take the well-known form

$$\frac{\partial \mathscr{H}}{\partial z^\alpha} - \frac{dp_\alpha}{dx} = 0, \quad \frac{\partial \mathscr{H}}{\partial p_\alpha} + \frac{dz^\alpha}{dx} = 0. \tag{5·6a}$$

Thus the function $\mathscr{H}$ defined by (5·3) appears as the most natural generalization of the Hamiltonian of mechanics, though other generalizations have also been given.*

Although $\mathscr{H}$ has been introduced as a component—X_1—of a vector in the natural co-ordinate system, it has an invariant significance, because for a given surface element of S the co-ordinate w normal to S has an invariant significance, and $\mathscr{H}$ is the conjugate to this co-ordinate w.

* The best known one is $\mathscr{H} = \mathscr{L} - p^i_\beta z^\beta_i$, first given by Volterra.

6. Boundary and initial data. The hyperbolic type

So far we have considered the extremal surfaces $z^\alpha(x)$ as given by their *boundary* $x^i(u)$, $z^\alpha(u)$ in the (x, z)-space, or in other words, by prescribing the "*boundary data*" $z^\alpha(u)$ on the whole boundary S in the x-space. Generally speaking, one understands by boundary data the prescribing of *one* quantity, e.g. z^α *or* P_α, on the whole of S.

For many considerations, however, in particular for deriving Poincaré's integral invariant, Lagrange brackets and Poisson brackets, it is necessary to characterize extremals by "*initial data*". Here we understand by initial data that on a part of the boundary S *two* data, e.g. z^α *and* P_α, are prescribed, while on another part of S no datum is prescribed.

In mechanics an extremal may in general be given by initial data equally well as by boundary data. In either case it is uniquely determined by its data* and depends continuously on them; the same is therefore true of the action integral I. There is thus no difficulty in passing over from the one standpoint to the other, as is done in the Hamilton-Jacobi theory of mechanics.

In a continuum theory the position is more complicated. One can distinguish between different types of systems according to what kind of data are "*correctly set*". This term has been explained in the introduction. There is one type of systems, the so-called "*elliptic type*", for which boundary data are in general correctly set, but initial data are not. In continuum physics all *equilibrium problems* belong to this type.

All physical problems of continuum *dynamics* belong to another type, called the "*hyperbolic type*". Before we can state precisely what kind of data are correctly set for this type, we have to give some preliminary explanations. For a detailed treatment we refer to the works of Hadamard and Courant, quoted in the introduction. There exists for the hyperbolic type a *real* $(n-1)$-dimensional "*characteristic cone*" (in physics called "light cone") at each point of the x-space, which gives rise to a classification of a given closed boundary S into different regions. For $n>2$ there are regions on S, called *space-like*, where the characteristic cone at a given point P of S intersects S only in P itself and in no other point of a suitable neighbourhood of P; there are other regions on S, called *time-like*, where the cone at P intersects S in an infinity of points in any given neighbourhood of P. The frontier between space-like and time-like regions on S is an $(n-2)$-dimensional manifold of points whose characteristic cones *touch* S. For $n=2$,

* The case where this is not true, e.g. "conjugate points", arises from an exceptional choice of data.

there are still the same different regions on S, but their roles of being space- or time-like may be interchanged, the characteristic cone degenerating into a pair of lines passing through P.

In physics, the history of a domain of space during an interval of time is described by a region D in space-time ($n=4$) of cylindrical shape. The boundary S of D consists of two bases, which are space-like, and of one curved part representing the history of the boundary of the domain of space, which is time-like. The principle of general relativity permits all those transformations of space-time which preserve the character of a region of being space- or time-like. Any closed surface S consisting of two distinct space-like regions, S_1 and S_2 say, and of a time-like region, joining S_1 with S_2, is relativistically equivalent to a cylinder of the kind described above; it is therefore admissible as a boundary for a problem of continuum dynamics.

For all problems of the hyperbolic type boundary data throughout S are in general not correctly set. The only correctly set data are: *initial data,* z^α *and* P_α*, on one of the two space-like regions of* S, S_1 *say; boundary data,* z^α *or* P_α*, on the time-like region; no data on the remaining space-like region,* S_2 *say.* Such data are called "*mixed data*".

For $n>3$, there exist still other types of systems, called "ultrahyperbolic", for which neither boundary data nor mixed data are in general correctly set. No physical problem belongs to such a type.

From now on we shall deal with the hyperbolic type only.

7. Interpretation of the boundary formula for the hyperbolic type

It is clear from the preceding considerations that the notions of integral invariant, Lagrange brackets and Poisson brackets can be extended only to continuous systems of the hyperbolic type and only as far as the initial data, i.e. the data on a space-like region are concerned. This, however, is precisely what physics requires.

On the other hand, the boundary formula (4·1) has only a formal meaning in the case of continuum dynamics, since the functional I does not depend uniquely and continuously on the boundary data. The conclusion which we have drawn from it, i.e. which quantities are pairs of conjugate variables, is also of a merely formal kind and is therefore valid for any type of system. But for our further considerations we have to adapt the boundary formula to the fact that I is now a proper functional of the mixed data and not of the boundary data on S.

It would be difficult to obtain an expression for the general functional differential of I, describing the variation of I corresponding to the variations of the mixed data. For our purpose, however, we need only study the special case where the boundary data on the time-like region are kept fixed and the initial data on the space-like region S_1 alone are varied. The boundary formula may then be written as

$$\delta \mathsf{I} = \left| \int^{(n-1)} (X_i \delta x^i + P_\alpha \delta z^\alpha) \, du \right|_{S_1}^{S_2}, \tag{7·1}$$

where we have inverted the sign of the variations on S_1. This formula should be compared with the corresponding formula (4·1 a) for the case of mechanics.

In (7·1), the variations on S_2 are determined by the initial data on S_1 and their variations. The right-hand side of (7·1) is, therefore, not a proper differential of the functional I considered as a functional of the initial data. In the following section we shall turn this fact to our advantage.

The fixed boundary data on the time-like region represent spatial boundary conditions holding throughout time. Their physical significance depends on the particular problem under consideration. For vibrations of a string or of a membrane they are conditions imposed on the boundary of the string or membrane (as is obvious). For a pure radiation field of finite extension, enclosed between walls, they represent the optical properties of the walls. For a radiation field extending through the whole of space, they represent conditions for the behaviour of the field at infinity in space. If the field contains charged particles, forming singular time-like world-lines extending in time (more precisely: proper-time) from the infinite past to the infinite future, one may exclude these singular lines by small tubes. One then obtains a multiply connected space-time, for which the boundary data on the time-like region consist not only of conditions at infinity but also of conditions on the surfaces of the small tubes. These latter conditions describe the behaviour of the singularities.

We need not deal with these conditions in the present work, since they are not of the nature of initial conditions and thus do not affect the Hamilton-Jacobi theory and quantization of a dynamical continuum, considered by itself. They may be of importance if one considers interaction processes, for instance the interaction of a radiation field with charged particles.

8. Integral invariant and Lagrange brackets

Let us consider a one-parametric *closed* family of extremals, labelled by a parameter λ, $0 \leqq \lambda \leqq 1$, such that the time-like parts of the boundaries are the same for the whole family. The initial data on S_1 then become functions of λ,

$$S_1: x^i = x^i(u, \lambda), \quad z^\alpha = z^\alpha(u, \lambda), \quad P_\alpha = P_\alpha(u, \lambda). \tag{8·1}$$

The fact that the family is closed is expressed by the relations

$$x^i(u, 0) = x^i(u, 1), \quad z^\alpha(u, 0) = z^\alpha(u, 1), \quad P_\alpha(u, 0) = P_\alpha(u, 1).$$

We thus obtain a tube of extremals in the (x, z)-space, which is completely determined by the tube of $(n-1)$-dimensional surfaces (8·1) and by the constant time-like boundary part. The functional I thus becomes a function of the parameter λ.

Let us now apply the modified boundary formula (7·1) to this tube by putting

$$\delta x^i = \frac{dx^i}{d\lambda} d\lambda, \quad \delta z^\alpha = \frac{dz^\alpha}{d\lambda} d\lambda$$

and then integrate it round the tube, i.e. from 0 to 1. The result must be zero since

$$\mathsf{I}(0) = \mathsf{I}(1).$$

Hence we obtain

$$\int_0^1 d\lambda \int_{S_1(\lambda)}^{(n-1)} \left(X_i \frac{dx^i}{d\lambda} + P_\alpha \frac{dz^\alpha}{d\lambda} \right) du = \int_0^1 d\lambda \int_{S_2(\lambda)}^{(n-1)} \left(X_i \frac{dx^i}{d\lambda} + P_\alpha \frac{dz^\alpha}{d\lambda} \right) du. \tag{8·2}$$

The formula (8·2) holds for any two tubes of $(n-1)$-dimensional space-like surfaces $S_1(\lambda)$ and $S_2(\lambda)$, provided that they both lie on the same tube of extremals. Hence the expression in (8·2) is invariant under a displacement along the tube. We have thus obtained a generalization of the notion of "*relative integral invariant*".*

$$\int_0^1 d\lambda \int_{S_1(\lambda)}^{(n-1)} \left(X_i \frac{dx^i}{d\lambda} + P_\alpha \frac{dz^\alpha}{d\lambda} \right) du. \tag{8·3}$$

* Cf. Cartan (1922, pp. 4, 9), Whittaker (1927, p. 272). Recently Born (1934) has attempted to derive a quantization method from Hilbert's "*independent integral*" (Hilbert 1900, 1905). Hilbert's integral can be given the same form as that of (8·3), but it applies to *selections* (Hilbert calls them "fields") of extremals such that through each point in the (x, z)-space there passes one and only one extremal. Hilbert's "*independence theorem*" is an immediate consequence of the invariance of (8·3). For if a field of extremals is especially chosen in such a way that the closed integral (8·3) vanishes for all tubes $S_1(\lambda)$ on a cross-section of the (x, z)-space, it must, owing to its invariance, do so throughout the field. This is Hilbert's theorem.

Since a displacement within the same extremal gives only a trivial contribution, we may without loss of generality assume that $\frac{dx^i}{d\lambda}=0$; then we obtain a generalization of *Poincaré's relative integral invariant*,

$$\int_0^1 d\lambda \int_{S_1(\lambda)}^{(n-1)} P_\alpha \frac{dz^\alpha}{d\lambda} du. \tag{8·4}$$

The P_α and z^α appearing under the integral sign in (8·4) are the initial data which determine the extremal if the boundary data on the time-like part of S are fixed.

The relative integral invariant (8·3) or (8·4) can be transformed into an *absolute integral invariant* with the help of Stokes' theorem. Consider a two-parametric family of extremals, labelled by the parameters ω_1, ω_2 such that the closed one-parametric family forms its boundary. It may be pictured as a "solid" tube of extremals in the (x, z)-space with the given tube as its surface. We then obtain for the integral invariants (8·3) and (8·4) respectively, with an obvious notation,

$$\iint d\omega_1 d\omega_2 \int_{S_1(\omega)}^{(n-1)} \left(\frac{d(X_i, x^i)}{d(\omega_1, \omega_2)} + \frac{d(P_\alpha, z^\alpha)}{d(\omega_1, \omega_2)}\right) du, \tag{8·5}$$

$$\iint d\omega_1 d\omega_2 \int_{S_1(\omega)}^{(n-1)} \frac{d(P_\alpha, z^\alpha)}{d(\omega_1, \omega_2)} du. \tag{8·6}$$

This is the generalization of Poincaré's absolute integral invariant (cp. Cartan 1922, pp. 19–20; Whittaker 1927, p. 272).

Again, without loss of generality we may work with the expression (8·6). As the domain of integration in the parameter space is arbitrary, the integrand under the integral-sign over ω must itself remain invariant under a displacement along the extremals. This integrand is a *generalized Lagrange bracket* (it will be referred to as L.B.) (cp. Whittaker 1927, p. 298),

$$\{\omega_1, \omega_2\} = \int_{S_1}^{(n-1)} \frac{d\{P_\alpha(u), z^\alpha(u)\}}{d(\omega_1, \omega_2)} du. \tag{8·7}$$

In the case $n=1$ the space-like region S_1 consists of a single point so that (8·7) reduces to the ordinary L.B., involving no integration but only the summation over α. In the general case we have the integration over S_1 besides the summation over the number of dependent variables.

In all the subsequent considerations the variables (u^r) will appear on the same footing as the index α. It is therefore useful to consider a point (u) of S_1 as a *continuous index*, which together with the discrete index α serves to label the dependent variables z and their conjugates P. From this standpoint the only difference between the dynamics of point systems and that of continua is the larger *magnitude* of indices which in the latter case is necessary in order to specify the initial data.

This standpoint also approaches the standpoint of Heisenberg and Pauli, who treated the continuum as the limiting case of a very large number of points.

In mechanics, all transformations preserving the L.B.'s are defined as "*canonical transformations*". The same definition will be adopted for the general case. Then the invariance property of the L.B.'s may be expressed by saying that for any dynamical continuum with fixed spatial boundary conditions the motion consists of a continuous unfolding of canonical transformations, starting from some initial data on some space-like region S_1.

9. Poisson brackets

In the considerations of the preceding section we have treated the P_α and z^α as functions of the u^r on S_1 and of the parameters ω_1, ω_2. We shall now invert the connexion and treat the quantities ω_1, ω_2 as functionals of the $P_\alpha(u)$, $z^\alpha(u)$, defined on S_1. The functional differential of such a functional ω has the form

$$\delta\omega = \int_{S_1}^{(n-1)} \left(\frac{\delta\omega}{\delta P_\alpha(u)} \delta P_\alpha(u) + \frac{\delta\omega}{\delta z^\alpha(u)} \delta z^\alpha(u) \right) du, \tag{9·1}$$

where $\dfrac{\delta\omega}{\delta P_\alpha(u)}$ and $\dfrac{\delta\omega}{\delta z^\alpha(u)}$ denote the functional derivatives of ω at the point (u).

In particular, we shall have to consider the functionals $P_\beta(v)$ and $z^\beta(v)$, i.e. the values of P_β and z^β at some given point (v) on S_1. We obtain from (9·1)

$$\left.\begin{aligned} \frac{\delta P_\beta(v)}{\delta P_\alpha(u)} &= \delta^\alpha_\beta\, \delta(v-u), & \frac{\delta P_\beta(v)}{\delta z^\alpha(u)} &= 0, \\ \frac{\delta z^\beta(v)}{\delta P_\alpha(u)} &= 0, & \frac{\delta z^\beta(v)}{\delta z^\alpha(u)} &= \delta^\beta_\alpha\, \delta(v-u), \end{aligned}\right\} \tag{9·2}$$

where $\delta(v-u)$ denotes the product of the $n-1$ δ-functions $\delta(v^r-u^r)$ of Dirac.

For any two functionals ω_1, ω_2 we define the *generalized Poisson bracket* (referred to as P.B.) (cp. Whittaker 1927, p. 299) by

$$[\omega_1, \omega_2] = \int_{S_1}^{(n-1)} \frac{d(\omega_1, \omega_2)}{d\{P_\alpha(u), z^\alpha(u)\}} du, \tag{9·3}$$

where the Jacobian under the integral sign is defined as

$$\frac{d(\omega_1, \omega_2)}{d\{P_\alpha(u), z^\alpha(u)\}} = \frac{\delta\omega_1}{\delta P_\alpha(u)}\frac{\delta\omega_2}{\delta z^\alpha(u)} - \frac{\delta\omega_1}{\delta z^\alpha(u)}\frac{\delta\omega_2}{\delta P_\alpha(u)}. \tag{9·4}$$

In the case of mechanics the generalized P.B.'s defined by (9·3) reduce to the ordinary P.B.'s. They also obey the same formal rules as the ordinary P.B.'s.

The P.B.'s have the same invariance properties as the L.B.'s. We shall show this by proving a relationship which allows us to express the P.B.'s in terms of the L.B.'s and vice versa. The prescribing of the initial data $P_\alpha(u)$, $z^\alpha(u)$ on S_1 (always for *fixed* boundary data on the time-like region of S) may be considered as a way of labelling all extremals (with the given fixed boundary data). There is, however, no need to choose especially the initial data for labelling the extremals. Any other set of the same magnitude may be used equally well. Let $\omega_a(v)$ $(a=1, 2, \dots, 2\nu;\ v \in S_1)$ be such a set which labels the extremals. We shall then have a one-one-correspondence

$$P_1(u) \dots P_\nu(u), \quad z^1(u) \dots z^\nu(u) \leftrightarrow \omega_1(v) \dots \omega_{2\nu}(v), \tag{9·5}$$

which may be interpreted as a *non-singular functional transformation* between two different modes of labelling the extremals.

Taking any two quantities, $\omega_a(v)$ and $\omega_b(v')$ say, we can form their L.B. and their P.B. The general relationship which we want to prove reads

$$\int_{S_1}^{(n-1)} dv \sum_{a=1}^{a=2\nu} \{\omega_a(v), \omega_b(v')\}\,[\omega_a(v), \omega_c(v'')] = \delta_c^b\,\delta(v'-v'') \quad (b, c = 1, 2, \dots, 2\nu).* \tag{9·6}$$

This relation expresses that the matrix formed by all the L.B.'s and the matrix formed by all the P.B.'s are contragredient (i.e. inverse and transposed) to each other. A corresponding relation holds in mechanics, with the only difference that we have here the continuous index v besides the discrete index a which results in the matrices having a continuous range of rows and columns. The proof of (9·6) therefore proceeds on the same lines as the corresponding proof in mechanics (Whittaker 1927, pp. 299–300).

To prove (9·6), we first use the definitions (8·7) and (9·3), (9·4) for the L.B.'s and P.B.'s so that (9·6) becomes

$$\int dv \sum_{a=1}^{a=2\nu} \int du \left(\frac{\delta P_\alpha(u)}{\delta\omega_a(v)}\frac{\delta z^\alpha(u)}{\delta\omega_b(v')} - \frac{\delta z^\alpha(u)}{\delta\omega_a(v)}\frac{\delta P_\alpha(u)}{\delta\omega_b(v')}\right) \times \int du' \left(\frac{\delta\omega_a(v)}{\delta P_\beta(u')}\frac{\delta\omega_c(v'')}{\delta z^\beta(u')} - \frac{\delta\omega_a(v)}{\delta z^\beta(u')}\frac{\delta\omega_c(v'')}{\delta P_\beta(u')}\right),$$

* It will be seen from the definition (8·7) of the L.B.'s that $\{\omega_a, \omega_b\}$ is contravariant in a and b. Since this fact is not expressed in our notation, we use the Σ-symbol for the summation over a.

it being understood that the integration always extends over the $(n-1)$-dimensional region S_1. In evaluating this expression we use the relations

$$\int dv \sum_{a=1}^{a=2\nu} \frac{\delta P_\alpha(u)}{\delta \omega_a(v)} \frac{\delta \omega_a(v)}{\delta P_\beta(u')} = \delta_\alpha^\beta \delta(u-u'), \quad \int dv \sum_{a=1}^{a=2\nu} \frac{\delta P_\alpha(u)}{\delta \omega_a(v)} \frac{\delta \omega_a(v)}{\delta z^\beta(u')} = 0,$$
$$\int dv \sum_{a=1}^{a=2\nu} \frac{\delta z^\alpha(u)}{\delta \omega_a(v)} \frac{\delta \omega_a(v)}{\delta P_\beta(u')} = 0, \qquad \int dv \sum_{a=1}^{a=2\nu} \frac{\delta z^\alpha(u)}{\delta \omega_a(v)} \frac{\delta \omega_a(v)}{\delta z^\beta(u')} = \delta_\beta^\alpha \delta(u-u'),$$

which follow directly from (9·2). We then obtain for the above expression

$$\int du \int du' \left(\frac{\delta z^\alpha(u)}{\delta \omega_b(v')} \frac{\delta \omega_c(v'')}{\delta z^\beta(u')} \delta_\alpha^\beta \delta(u-u') + \frac{\delta P_\alpha(u)}{\delta \omega_b(v')} \frac{\delta \omega_c(v'')}{\delta P_\beta(u')} \delta_\beta^\alpha \delta(u-u') \right)$$
$$= \int du \left(\frac{\delta z^\alpha(u)}{\delta \omega_b(v')} \frac{\delta \omega_c(v'')}{\delta z^\alpha(u)} + \frac{\delta P_\alpha(u)}{\delta \omega_b(v')} \frac{\delta \omega_c(v'')}{\delta P_\alpha(u)} \right) = \delta_c^b \delta(v'-v''),$$

so that (9·6) is proved.

It follows that the P.B.'s are invariant under canonical transformations.

10. Classical laws in terms of Poisson brackets

As in mechanics, we can formulate explicit conditions which are necessary and sufficient for a transformation of the $P_\alpha(u)$, $z^\alpha(u)$ to be canonical. For, from the formulae (9·2)–(9·4), we obtain for trivial reasons

$$\left.\begin{aligned} [P_\alpha(v), P_\beta(v')] = 0, \quad [z^\alpha(v), z^\beta(v')] = 0, \\ [P_\alpha(v), z^\beta(v')] = \delta_\alpha^\beta \delta(v-v'), \end{aligned}\right\} \tag{10·1}$$

and it follows from the invariance of the P.B.'s that the transformed quantities, if obtained by a canonical transformation, must satisfy the same relations. This proves that the relations (10·1) are necessary.

By differentiating the second and the third relation of (10·1) with respect to v'^r we obtain, taking account of (5·1),

$$[z^\alpha(v), \tilde{z}_r^\beta(v')] = 0, \quad [P_\alpha(v), \tilde{z}_r^\beta(v')] = -\delta_\alpha^\beta \delta_r(v-v'), \tag{10·2}$$

where $\delta_r(v-v')$ stands for $\frac{\mathfrak{d}}{\mathfrak{d} u^r}\delta(v-v')$. Now let $\mathscr{F}(P_\alpha, z^\alpha, \tilde{z}_r^\alpha)$ be an arbitrary function defined on S_1 and put

$$\mathsf{F} = \int_{S_1}^{(n-1)} \mathscr{F}(v)\, dv, \quad (v \in S_1). \tag{10·3}$$

We can evaluate the P.B.'s of $P_\alpha(u)$ and $z^\alpha(u)$ with F with the help of (10·1) and (10·2) and obtain

$$[P_\alpha(u), \mathsf{F}] = \frac{\partial \mathscr{F}}{\partial z^\alpha} - \frac{\mathfrak{d}}{\mathfrak{d} u^s} \frac{\partial \mathscr{F}}{\partial \tilde{z}_s^\alpha}, \quad [z^\alpha(u), \mathsf{F}] = -\frac{\partial \mathscr{F}}{\partial P_\alpha}. \tag{10·4}$$

In this manner we can evaluate any P.B.'s, using solely the relations (10·1), the relations (10·2) being a consequence of (10·1). It follows, therefore, that (10·1) is also sufficient for a transformation to be canonical.

We shall now apply the general formulae (10·4) to the quantities $\mathscr{H}$ and h_r, defined in section 5. We put

$$\mathsf{H} = \int_{S_1}^{(n-1)} \mathscr{H}(v)\,dv, \quad \mathsf{h}_r = \int_{S_1}^{(n-1)} h_r(v)\,dv. \tag{10·5}$$

Applying (10·4) to the case $\mathscr{F} = h_r$, we obtain from the last set of (5·3)

$$[P_\alpha(u), \mathsf{h}_r] = \frac{\mathfrak{d}P_\alpha}{\mathfrak{d}u^r}, \quad [z^\alpha(u), \mathsf{h}_r] = \frac{\mathfrak{d}z^\alpha}{\mathfrak{d}u^r}. \tag{10·6}$$

Applying (10·4) to the case $\mathscr{F} = \mathscr{H}$ and using the canonical equations (5·6), we obtain

$$[P_\alpha(u), \mathsf{H}] = \frac{\mathfrak{d}P_\alpha}{\mathfrak{d}w}, \quad [z^\alpha(u), \mathsf{H}] = \frac{\mathfrak{d}z^\alpha}{\mathfrak{d}w}. \tag{10·7}$$

Since S_1 may be any space-like region, w may be any time-like co-ordinate.

Conversely, the canonical equations (5·6) follow from (10·7) with the help of the general formula (10·4). The equations (10·7) may therefore be considered as the formulation of the canonical equations by means of P.B.'s. In the case of mechanics they go over into the well-known formulae

$$[p_\alpha, \mathscr{H}] = \frac{dp_\alpha}{dx}, \quad [z^\alpha, \mathscr{H}] = \frac{dz^\alpha}{dx}. \tag{10·7a}$$

In the case $n > 1$ they are supplemented by the equations (10·6). (10·7) and (10·6) form together a set of covariant vector equations in the x-space.

It follows from (10·6) and (10·7) that we have for an arbitrary function $\mathscr{F}(P_\alpha, z^\alpha, \tilde{z}^\alpha_r)$,

$$[\mathscr{F}(u), \mathsf{H}] = \frac{\mathfrak{d}\mathscr{F}}{\mathfrak{d}w}, \quad [\mathscr{F}(u), \mathsf{h}_r] = \frac{\mathfrak{d}\mathscr{F}}{\mathfrak{d}u^r}. \tag{10·8}$$

In this way every dynamical relation of the continuum can be expressed in terms of the P.B.'s.

11. Quantization

Having formulated classical continuum dynamics in terms of Poisson brackets, the quantization can be performed in the same way as in mechanics. The continuous set of initial data $P_\alpha(u)$, $z^\alpha(u)$ becomes a continuous set of non-commuting observables, and a general argument, which is valid irrespective of whether the set is discrete or continuous, requires that the quantum P.B.'s, which are defined by the property of obeying the same formal rules as the classical P.B.'s, must be proportional to the commutation brackets. In this way we obtain ultimately (Dirac 1935, pp. 89–90)

$$[\omega_1, \omega_2] = \frac{i}{\hbar}(\omega_1\omega_2 - \omega_2\omega_1). \tag{11·1}$$

With this new interpretation of the P.B.'s the relations (10·1) and 10·2) become the "*Quantum Conditions*", the function $\mathscr{H}(P_\alpha, z^\alpha, z^\alpha_r)$ or better its integral H becomes the "*Quantum Hamiltonian*", and the relations (10·7) become the "*Quantum Equations of Motion*" of the quantized continuum. The Hamiltonian has an invariant significance, as we have seen at the end of section 5. The relations (10·4) and (10·6) follow from (10·1) and (10·2), and (10·8) follows from (10·6) and (10·7) by the same arguments as in the classical theory.

The quantum formalism thus obtained is the same as the one we had obtained in I (Weiss 1936, pp. 201–4) by an argument based on a comparison with the results of Heisenberg and Pauli. The present formalism applies to any space-like region S_1 and the corresponding time-like co-ordinate w normal to S_1, but it does not and, on physical grounds, should not apply to a time-like region.*

Summary

The Hamilton-Jacobi theory of point mechanics is extended to the mechanics of continuous media, following on the lines first proposed by Prange. The method is based on the equivalence between the Euler equations and the "boundary formula" (*formule aux limites*) in the calculus of variations. It is shown that the notions of Lagrange brackets and Poisson brackets can be extended, but only if the Euler equations of the continuum are of hyperbolic type. Consequently, these notions only apply to a dynamical continuum and not to equilibrium problems. Except for this restriction, the method is applicable quite generally, for linear as well as for non-linear theories. Once the dynamical laws are expressed in terms of Poisson brackets, the transition to the quantum theory can be effected by a brief, formal argument in the same way as in point mechanics.

* At the end of I, I tried to apply the quantization to a time-like region as well. This was due to my not having sufficiently taken account of the hyperbolic character of all dynamical problems. No quantum relations on a time-like region are justified.

REFERENCES

Born 1934 *Proc. Roy. Soc.* A, **143**, 410.
Cartan 1922 "Leçons sur les invariants intégraux." Paris.
Courant-Hilbert 1937 "Methoden der mathematischen Physik", **2**. Berlin.
De Donder 1935 "Théorie invariantive du calcul des variations." Paris.
Dirac 1935 "The principles of quantum mechanics", 2nd ed. Oxford.
Hadamard 1910 "Leçons sur le calcul des variations." Paris.
— 1923 "Lectures on Cauchy's problem." Yale, Oxford.
— 1932 "Leçons sur le problème de Cauchy etc." Paris.
Heisenberg and Pauli 1929 *Z. Phys.* **56**, 1.
Hilbert 1900 *Nachr. Ges. Wiss. Göttingen*, p. 253.
— 1905 *Nachr. Ges. Wiss. Göttingen*, p. 159.
Juvet 1926 "Sur une équation aux dérivées fonctionnelles partielles etc." Thèse, Paris.
Prange 1915 "Die Hamilton-Jacobi'sche Theorie für Doppelintegrale." Dissertation, Göttingen.
Volterra 1913 "Leçons sur les fonctions de lignes." Paris.
— 1930 "Theory of functionals." London.
Volterra and Pérès 1937 "Théorie générale des fonctionnelles", **1**. Paris.
Weiss 1936 *Proc. Roy. Soc.* A, **156**, 192.
Whittaker 1927 "Analytical dynamics", 3rd ed. Cambridge.

On the Hamilton-Jacobi theory and quantization of generalized electrodynamics

BY P. WEISS, PH.D., *Downing College, Cambridge*

(*Communicated by R. H. Fowler, F.R.S.—Received* 1 *September* 1938)

1. INTRODUCTION

In the preceding paper, which will be quoted as A, the Hamilton-Jacobi theory has been developed for a dynamical continuum of quite general type, and it has been shown that its classical laws can be expressed in terms of Poisson brackets involving pairs of canonically conjugate variables. In this way the theory of a dynamical continuum can be treated on the same lines as the dynamical theory of point systems.

This result is of importance for the study of the procedure of quantization, for it is a well-known fact that a quantum Poisson bracket which obeys the same algebraic rules as a classical Poisson bracket must be proportional to a commutation bracket (Dirac 1935, p. 89). This fact is independent of

Chapter 23
Markus Fierz and Wolfgang Pauli (1939): On Relativistic Wave Equations for Particles of Arbitrary Spin in an Electromagnetic Field

Markus Fierz and Wolfgang Pauli (1939). On Relativistic Wave Equations for Particles of Arbitrary Spin in an Electromagnetic Field. *Proceedings of the Royal Society of London*, A173, 211–232.

Keesom, W. H. 1921 *Phys. Z.* **22**, 129.
Keyes, F. G. and co-workers 1922 *J. Math. Phys. Mass. Inst. Tech.* **1**, 191.
— 1927 *J. Amer. Chem. Soc.* **49**, 1403.
Lewis, D. T. 1938 *J. Chem. Soc.* pp. 261, 1056, 1061.
Mack, E. and Melaven, R. M. 1931 *J. Amer. Chem. Soc.* **54**, 888.
Sidgwick, N. V. 1932 *Ann. Rep. Chem. Soc.* p. 67.
— 1933 *The covalent link.* Cornell Univ. Press.
Whytlaw-Gray, R. and Burt, F. P. 1909 *J. Chem. Soc.* **95**, 1633.
Woodhead, M. and Whytlaw-Gray, R. 1933 *J. Chem. Soc.* p. 846.

On relativistic wave equations for particles of arbitrary spin in an electromagnetic field

By M. Fierz and W. Pauli
Physikalisches Institut der Eidgenössischen Technischen Hochschule, Zürich

(*Communicated by P. A. M. Dirac, F.R.S.—Received* 31 *May* 1939)

1. Introduction

The investigations of Dirac (1936) on relativistic wave equations for particles with arbitrary spin have recently been followed up by one of us (Fierz, 1939, referred to as (A)). It was there found possible to set up a scheme of second quantization in the absence of an external field, and to derive expressions for the current vector and the energy-momentum tensor. These considerations will be extended in the present paper to the case when there is an external electromagnetic field, but we shall in the first instance disregard the second quantization and confine ourselves to a c-number theory.

The difficulty of this problem is illustrated by the fact that the most immediate method of taking into account the effect of the electromagnetic field, proposed by Dirac (1936), leads to inconsistent equations as soon as the spin is greater than 1. To make this clear we consider Dirac's equations for a particle of spin 3/2, which in the force-free case run as follows:

$$\left.\begin{aligned} \kappa b_{\gamma}^{\dot{\alpha}\dot{\beta}} &= p^{\dot{\alpha}\rho} a_{\gamma\rho}^{\dot{\beta}} = p^{\dot{\beta}\rho} a_{\gamma\rho}^{\dot{\alpha}}, \\ \kappa a_{\alpha\beta}^{\dot{\gamma}} &= p_{\alpha\dot{\rho}} b_{\beta}^{\dot{\rho}\dot{\gamma}} = p_{\beta\dot{\rho}} b_{\alpha}^{\dot{\rho}\dot{\gamma}}, \end{aligned}\right\} \tag{1}$$

where $a^{\dot\gamma}_{\alpha\beta} = a^{\dot\gamma}_{\beta\alpha}$ and $b^{\dot\alpha\dot\beta}_{\gamma} = b^{\dot\beta\dot\alpha}_{\gamma}$ are symmetrical spinors. Dirac attempted to take the external electromagnetic field into account by replacing the spinor $p_{\alpha\dot\beta}$ by $\Pi_{\alpha\dot\beta}$, which arises from it by substituting $-i\partial/\partial x_k - e\phi_k/hc$ for $-i\partial/\partial x_k$ (ϕ_k being the electromagnetic potentials). The $\Pi_{\alpha\dot\beta}$ are then non-commuting operators satisfying the relations (cf. Appendix)

$$\Pi_{\dot\alpha\beta}\Pi^{\dot\gamma\delta} - \Pi^{\dot\gamma\delta}\Pi_{\dot\alpha\beta} = \delta_{\dot\alpha}{}^{\dot\gamma} f_{\beta}{}^{\delta} + \delta_{\beta}{}^{\delta} f_{\dot\alpha}{}^{\dot\gamma}, \tag{2·1}$$

where $f_{\alpha\beta}$ and $f_{\dot\alpha\dot\beta}$ are the two symmetrical spinors ($f_{\alpha}{}^{\alpha} = f_{\dot\alpha}{}^{\dot\alpha} = 0$) associated with the antisymmetrical field tensor f_{ik}. By contracting this relation it follows that

$$\Pi_{\alpha\dot\rho}\Pi^{\dot\rho\beta} - \Pi^{\dot\rho\beta}\Pi_{\alpha\dot\rho} = 2f_{\alpha}{}^{\beta}.$$

On the other hand, just as in the force-free case, we have

$$\Pi_{\alpha\dot\rho}\Pi^{\dot\rho\beta} + \Pi^{\dot\rho\beta}\Pi_{\alpha\dot\rho} = -2\Pi^2\delta_{\alpha}{}^{\beta}, \tag{2·2}$$

and so

$$\Pi_{\alpha\dot\rho}\Pi^{\dot\rho\beta} = -\Pi^2\delta_{\alpha}{}^{\beta} + f_{\alpha}{}^{\beta}. \tag{2·3}$$

From Dirac's proposed equations

$$\kappa b^{\dot\alpha\dot\beta}_{\gamma} = \Pi^{\dot\alpha\rho} a^{\dot\beta}_{\gamma\rho} = \Pi^{\dot\beta\rho} a^{\dot\alpha}_{\gamma\rho}, \tag{3·1}$$

$$\kappa a^{\dot\gamma}_{\alpha\beta} = \Pi_{\alpha\dot\rho} b^{\dot\gamma\dot\rho}_{\beta} = \Pi_{\beta\dot\rho} b^{\dot\gamma\dot\rho}_{\alpha}, \tag{3·2}$$

however, it would follow that

$$\Pi_{\alpha\dot\rho}\Pi^{\dot\rho\sigma} a^{\dot\gamma}_{\beta\sigma} = \kappa\Pi_{\alpha\dot\rho} b^{\dot\rho\dot\gamma}_{\beta} = \kappa^2 a^{\dot\gamma}_{\alpha\beta},$$

and so from (2·3)

$$-\Pi^2 a^{\dot\gamma}_{\alpha\beta} + f_{\alpha}{}^{\sigma} a^{\dot\gamma}_{\beta\sigma} = \kappa^2 a^{\dot\gamma}_{\alpha\beta}.$$

Since the right-hand side is symmetrical in α and β it follows that the subsidiary condition

$$f_{\alpha}{}^{\sigma} a^{\dot\gamma}_{\beta\sigma} = f_{\beta}{}^{\sigma} a^{\dot\gamma}_{\alpha\sigma},$$

or

$$f^{\rho\sigma} a^{\dot\gamma}_{\rho\sigma} = 0, \tag{3·3}$$

must be satisfied by the spinor field $a^{\dot\gamma}_{\alpha\beta}$; but this cannot in general be satisfied simultaneously with the other equations.

One might at first hope to avoid this objection by replacing (3·1), (3·2) by the weaker conditions

$$2\kappa b^{\dot\alpha\dot\beta}_{\gamma} = \Pi^{\dot\alpha\rho} a^{\dot\beta}_{\gamma\rho} + \Pi^{\dot\beta\rho} a^{\dot\alpha}_{\gamma\rho}, \tag{4·1}$$

$$2\kappa a^{\dot\gamma}_{\alpha\beta} = \Pi_{\alpha\dot\rho} b^{\dot\gamma\dot\rho}_{\beta} + \Pi_{\beta\dot\rho} b^{\dot\gamma\dot\rho}_{\alpha}. \tag{4·2}$$

It is to be remarked, however, that even in the force-free case such a system no longer leads to a wave equation of second order. (A closer discussion shows that these equations describe, besides particles of spin 3/2 and rest-mass κ, also particles of spin 1/2 and rest-mass 2κ). And further, the expression for the total charge turns out to be no longer positive definite, and this makes quantization consistent with the exclusion principle impossible (for quantization consistent with Bose statistics the total energy is, on the other hand, not positive for the case of half-integral spins).

This modification was therefore abandoned and the equations (1) were retained for the force-free case. The problem then arose, besides replacing the $p_{\alpha\dot{\beta}}$ by the $\Pi_{\alpha\dot{\beta}}$, of adding to these equations extra terms, depending on the field strengths, in such a way that they remained self-consistent in the presence of an external field.

A completely analogous problem arises for integral spins. For instance, the field equations for spin 2, involving according to (A) a symmetrical tensor A_{ik}, whose trace $\sum\limits_i A_{ii}$ vanishes, are

$$\Box A_{ik} = \kappa^2 A_{ik}, \tag{5·1}$$

$$\frac{\partial A_{ik}}{\partial x_i} = 0. \tag{5·2}$$

The second set of conditions is indispensable if the total energy is to be positive definite. In fact if they were omitted those waves with only components of the type A_{4i} would give rise to negative values of the total energy. On the other hand, the equations which arise from (5·1), (5·2) when $\partial/\partial x_k$ is replaced by $\partial/\partial x_k - ie\phi_k/\hbar c$ are not compatible, for the operators $\Pi^2 \equiv \sum\limits_k \Pi_k^2$ and Π_k are not commutative ($\Pi_k = -i\partial/\partial x_k - e\phi_k/\hbar c$).

We shall not attack the problem of deriving such additional terms to make the equations compatible directly but solve it by an artifice. This consists in introducing auxiliary tensors or spinors of lower rank than the original ones (for spin 3/2 they will be simple spinors c_α and $d_{\dot{\alpha}}$; for spin 2 a scalar C) and deriving all equations from a variation principle without having to introduce extra conditions. By suitably choosing the numerical coefficients in the Lagrange function it will follow from the field equations (derived from the variation) that in the absence of an external field the auxiliary quantities vanish and the additional conditions (5·2) or (1) are satisfied automatically (cf. § 2, equations (10), (11)).

That such a procedure is reasonable seems to be shown by the fact that, for vanishing rest-mass, our equations for the case of spin 2 go over into those of the relativity theory of weak gravitational fields (i.e. $g_{\mu\nu} = \delta_{\mu\nu} + \gamma_{\mu\nu}$, neglecting terms of order higher than the first in $\gamma_{\mu\nu}$); the "gauge-transformations" are identical with the changes induced in $\gamma_{\mu\nu}$ by infinitesimal co-ordinate transformations (§ 6).

Although the following only deals in detail with the interaction of the particles and an external electromagnetic field, the interaction with other particles which can be absorbed and emitted could be formulated analogously. For instance, the interaction with a new scalar field ψ could be introduced by extra terms in the Lagrange function which arise from those in the Lagrangian of the following pages containing a factor κ^2 by replacement of κ^2 by ψ. On the other hand, it is important that a one-to-one correspondence should be possible between the states (eigenfunctions) with the external field and without. This is equivalent to saying that the number of conditions which the field and auxiliary variables (and their time-derivatives for integral spins) must satisfy *at a definite time* is not diminished by the presence of an external field. Otherwise, as is illustrated in Appendix I by a special example with particles of spin 1, singularities occur when the external field is made to vanish slowly. In the main text, however, this requirement of the continued existence of subsidiary conditions in an external field is always fulfilled.

This requirement also seems important for the second quantization of the fields, a topic not treated in detail here. It enables one, namely, starting from the commutation rules of (A), to make an expansion of the commutation brackets of all field quantities in powers of the charge e. It is to be remarked that with particles of spin greater than 1 the charge-densities at different points no longer commute. A closer study of this circumstance, which strongly distinguishes the spin values of 0, 1/2, 1 (cf. A, Introduction) is to be desired.

As may be seen from our last section (§ 8), our aim was not so much to set up the most general possible relativistic equations for particles of higher spin but rather to show that, in the present state of the theory, the existence of elementary particles of spin higher than 1 cannot be excluded, although the theory for such particles is considerably more complicated than for smaller spin values. In this connexion it may be mentioned that we have been unable to generalize the field-equations which in the notation of (A) correspond to $k \neq l$, or the current-vectors $s^{(q)}$ for which $q > 1$ (cf. (A), I, II, III and (5·6)).

I. SPIN 2

2. Derivation of the force-free equations from a variation principle

As an example of the theory of a wave-field corresponding to particles of spin higher than 1 in interaction with other fields let us first consider the theory for spin 2.

As was shown in (A) such a field is described in the absence of external fields by a symmetrical tensor A_{ik} of second rank, whose trace is zero, satisfying the wave equation

$$\Box A_{ik} = \kappa^2 A_{ik}, \tag{5·1}$$

and for which the additional condition

$$\frac{\partial A_{ik}}{\partial x_i} = 0 \tag{5·2}$$

is fulfilled. In this the indices i, k run from 1 to 4. (x_i) stands for (x, y, z, ict) and $\Box \equiv \nabla^2 - (1/c^2)\,\partial^2/\partial t^2$. Summation over indices occurring twice is to be understood. It can be shown that the total energy of the field is positive only if the extra condition (5·2) is satisfied, i.e. if the vector $\partial A_{ik}/\partial x_i$ does not vanish it describes particles of negative energy. If one introduces external fields, this must be done in such a manner that after they have been shut off the condition (5·2) is again fulfilled, so that no new particles of negative energy should be created.

In order to discover a correct generalization of equations (5·1), (5·2) for external forces we shall look for a variation principle

$$\delta \int L\,d\Omega = 0,$$

from which (5·1) and (5·2) can be derived. At this point it is useful to introduce an auxiliary scalar field C, on which L will be taken to depend and which is to be varied independently of A_{ik}. The introduction of C is an artifice which enables one to derive the additional condition (5·2) from the Lagrange function by variation. For simplicity let us assume that A_{ik} and C are "real" fields, i.e.

$$A^*_{ik} = A_{ik}; \quad C^* = C.$$

(Here the tensor $A^*_{ik\,\ldots\,l}$, conjugate to the tensor $A_{ik\,\ldots\,l}$, is equal to $(-)^n \bar{A}_{ik\,\ldots\,l}$, where n is the number of times 4 appears among the indices $i, k, \ldots, l$, and the bar denotes the complex conjugate. "Real" tensors are those for which $A^*_{ik\,\ldots\,l} = A_{ik\,\ldots\,l}$.)

For the function L we make the following choice:

$$L = \kappa^2 A_{ik}A_{ik} + \frac{\partial A_{ik}}{\partial x_l}\frac{\partial A_{ik}}{\partial x_l} + a_1 \frac{\partial A_{rk}}{\partial x_r}\frac{\partial A_{sk}}{\partial x_s} + a_2 \kappa^2 C^2 + a_3 \frac{\partial C}{\partial x_l}\frac{\partial C}{\partial x_l} + \frac{\partial A_{rk}}{\partial x_r}\frac{\partial C}{\partial x_k}. \quad (6)$$

A_{ik} is of course symmetrical and fulfils the trace condition $A_{ii} = 0$, a fact which must be remembered when performing the variation. By varying A_{ik} and C we obtain the following equations:

$$2\kappa^2 A_{ik} - 2\Box A_{ik} - a_1 \left\{ \frac{\partial^2 A_{rk}}{\partial x_r \partial x_i} + \frac{\partial^2 A_{ri}}{\partial x_r \partial x_k} - \tfrac{1}{2}\delta_{ik} \frac{\partial^2 A_{rs}}{\partial x_r \partial x_s} \right\} - \frac{\partial^2 C}{\partial x_i \partial x_k} + \tfrac{1}{4}\delta_{ik}\Box C = 0, \quad (7\cdot 1)$$

$$2a_2 \kappa^2 C - 2a_3 \Box C - \frac{\partial^2 A_{rk}}{\partial x_r \partial x_k} = 0. \quad (7\cdot 2)$$

Let us now determine the three constants a_1, a_2, a_3 in such a way that $\partial A_{ik}/\partial x_i$ and C vanish as a consequence of (7·1) and (7·2). For this we differentiate (7·1) with respect to x_i and obtain

$$2\kappa^2 \frac{\partial A_{ik}}{\partial x_i} = 2\Box \frac{\partial A_{ik}}{\partial x_i} + a_1 \left\{ \Box \frac{\partial A_{ik}}{\partial x_i} + \frac{1}{2} \frac{\partial^3 A_{si}}{\partial x_i \partial x_s \partial x_k} \right\} + \tfrac{3}{4}\Box \frac{\partial C}{\partial x_k}. \quad (8)$$

If we now put $a_1 = -2$, the right-hand side will only contain derivatives of the scalars C and $\partial^2 A_{ik}/\partial x_i \partial x_k$, which latter we shall denote by A for brevity:

$$\frac{\partial^2 A_{ik}}{\partial x_i \partial x_k} \equiv A.$$

The equation (8) then becomes

$$2\kappa^2 \frac{\partial A_{ik}}{\partial x_i} = -\frac{\partial A}{\partial x_k} + \tfrac{3}{4}\Box \frac{\partial C}{\partial x_k}. \quad (8')$$

This equation means that the vector $\partial A_{ik}/\partial x_i$ is the gradient of a scalar, and can therefore describe only particles of spin zero.

Now let us differentiate (8) with respect to x_k and obtain together with (7·2) the two equations

$$2\kappa^2 A + \Box A - \tfrac{3}{4}\Box\Box C = 0, \quad (9)$$

$$-A + 2a_2\kappa^2 C - 2a_3 \Box C = 0. \quad (7\cdot 2')$$

This is a linear, homogeneous system of equations for A and C. We now choose a_2 and a_3 in such a manner that the operator determinant of the system shall never vanish; A and C will then vanish, and thus according to (8) also $\partial A_{ik}/\partial x_i$.

One obtains the following expression for the determinant

$$4a_2\kappa^4 + 2\kappa^2(a_2 - 2a_3)\Box - (2a_3 + \tfrac{3}{4})\Box\Box. \tag{10}$$

If we put $$a_3 = -\tfrac{3}{8}, \quad 2a_3 = a_2 = -\tfrac{3}{4},$$

then the determinant has the value $-3\kappa^4$, and since κ is supposed to be different from zero, it will never be zero. We therefore have

$$A = 0, \quad C = 0, \quad \frac{\partial A_{ik}}{\partial x_i} = 0.$$

We are thus led to the Lagrange function

$$L = \kappa^2 A_{ik}A_{ik} + \frac{\partial A_{ik}}{\partial x_l}\frac{\partial A_{ik}}{\partial x_l} - 2\frac{\partial A_{rk}}{\partial x_r}\frac{\partial A_{sk}}{\partial x_s} - \tfrac{3}{4}\kappa^2 C^2 - \frac{3}{8}\frac{\partial C}{\partial x_l}\frac{\partial C}{\partial x_l} + \frac{\partial A_{rk}}{\partial x_r}\frac{\partial C}{\partial x_k}, \tag{11}$$

and the corresponding field equations

$$2\kappa^2 A_{ik} - 2\Box A_{ik} + 2\left\{\frac{\partial^2 A_{sk}}{\partial x_s \partial x_i} + \frac{\partial^2 A_{si}}{\partial x_s \partial x_k} - \tfrac{1}{2}\delta_{ik}\frac{\partial^2 A_{rs}}{\partial x_r \partial x_s}\right\} - \frac{\partial^2 C}{\partial x_i \partial x_k} + \tfrac{1}{4}\delta_{ik}\Box C = 0, \tag{12·1}$$

$$-\tfrac{3}{2}\kappa^2 C + \tfrac{3}{4}C - \frac{\partial^2 A_{rk}}{\partial x_r \partial x_k} = 0, \tag{12·2}$$

from which one can derive equations (5·1) and (5·2), and also the equation $C = 0$.

As we are interested in the influence of external forces on the field A_{ik} it will be useful to use a notation in which the time is separated from the other co-ordinates. We shall therefore discuss equations (12·1), (12·2) from this point of view.

The field in the example considered belongs to the spin value $f = 2$, and therefore gives $2f+1 = 5$ states for a given direction and frequency. The differential equations for the fields A_{ik} and C are of second order. At a given time, therefore, one can prescribe the values of 5 components of A_{ik} and their time-derivatives at all points of space. Since the field A_{ik} has altogether $(f+1)^2 = 9$ components, there remain, together with the one component of C, 5 components and their first time-derivatives which cannot be given at will. That is to say, there must be 10 subsidiary conditions,* containing

* We would like to point out that we use the term "additional conditions" in the sense of equations not following from a variation principle giving the main equations (§2, §4), whereas the term "subsidiary conditions" refers to equations derived from the variation but which have the effect of reducing the number of degrees of freedom.

perhaps higher space-derivatives, but only first derivatives with respect to time, from which, if 5 components of A_{ik} and their time-derivatives are given, one can calculate the remaining ten quantities. These conditions can be derived from equations (12·1) and (12·2).

In the above-defined sense the following equations are to be regarded as subsidiary conditions:

$$C = 0, \tag{13·1}$$

$$\frac{\partial C}{\partial x_4} = 0. \tag{13·2}$$

Let us further consider (12·1) for $i \neq 4$, $k = 4$. To be able to write the equation conveniently in this form we introduce Greek indices $\alpha, \beta, \ldots$, which run only from 1 to 3. We then get

$$2\kappa^2 A_{\alpha 4} - 2\frac{\partial^2}{\partial x_\beta^2} A_{\alpha 4} + 2\frac{\partial^2 A_{\beta 4}}{\partial x_\alpha \partial x_\beta} + 2\frac{\partial^2 A_{\alpha\beta}}{\partial x_\beta \partial x_4}$$
$$+ 2\frac{\partial^2 A_{44}}{\partial x_\alpha \partial x_4} - \frac{\partial^2 C}{\partial x_\alpha \partial x_4} = 0, \quad (\alpha = 1, 2, 3). \qquad (13·3)\ldots(13·5)$$

These constitute three more conditions, since the second time-derivatives of $A_{\alpha 4}$ have dropped out. By adding the (4, 4)-component of (12·1) and (12·2) we get a sixth condition:

$$2\kappa^2 A_{44} - 2\frac{\partial^2 A_{44}}{\partial x_\beta^2} - \frac{\partial^2 A_{\alpha\beta}}{\partial x_\alpha \partial x_\beta} + \frac{\partial^2 C}{\partial x_\beta^2} - \tfrac{3}{2}\kappa^2 C = 0. \tag{13·6}$$

Differentiating (12·1) with respect to x_i we obtain

$$2\kappa^2 \frac{\partial A_{ik}}{\partial x_i} + \frac{\partial^3 A_{is}}{\partial x_k \partial x_i \partial x_s} - \frac{3}{4}\frac{\partial}{\partial x_k}\Box C = 0.$$

Combining this with (12·2) we obtain

$$2\frac{\partial A_{ik}}{\partial x_i} - \frac{3}{2}\frac{\partial C}{\partial x_k} = 0. \qquad (13·7)\ldots(13·10)$$

The above equation holds for $k = 1, \ldots, 4$. We have therefore found 10 subsidiary conditions.

3. Introduction of interactions

The theory as presented up till now is equivalent to the theory in (A). By adding suitable terms to the Lagrangian we can introduce interactions with other fields. One must take care, however, that the subsidiary conditions are not impaired. This would mean that the dimensionality of the manifold of states was altered by "switching on" the forces, and it turns out that these new states give rise to singularities when the field is "switched off", as we shall illustrate by means of an example in the Appendix. We shall here consider the effect of an electromagnetic field.

In this case we must naturally assume that the fields A_{ik} and C are complex. Let ϕ_k be the four-potential of the electromagnetic field, e the charge of the particles. Then we have

$$i\Pi_k = \frac{\partial}{\partial x_k} - \frac{ie}{hc}\phi_k,$$

and
$$f_{ik} \equiv \Pi_i \Pi_k - \Pi_k \Pi_i = \frac{ie}{hc}\left\{\frac{\partial \phi_k}{\partial x_i} - \frac{\partial \phi_i}{\partial x_k}\right\}.$$

We take the following form for L:

$$L = \kappa^2 A^*_{ik} A_{ik} + \Pi_l A_{ik} \Pi^*_l A^*_{ik} - 2\Pi_r A_{rk} \Pi^*_s A^*_{sk} + f_{ir} A_{rk} A^*_{ik} + \tfrac{1}{2}\{\Pi_r A_{rk} \Pi^*_k C^* + \Pi^*_r A^*_{rk} \Pi_k C\} - \tfrac{3}{4}\kappa^2 C^* C - \tfrac{3}{8}\Pi_l C \Pi^*_l C^*. \quad (14)$$

The term $f_{ir} A_{rk} A^*_{ik}$ (proportional to the field strengths) has been added because then the derivation of the subsidiary conditions is particularly simple, but it is not necessary. The Lagrange function can also be brought into another form by partial integration, namely,

$$L = \kappa^2 A^*_{ik} A_{ik} + \tfrac{1}{2}\{\Pi_l A_{ik} - \Pi_i A_{lk}\}\{\Pi^*_l A^*_{ik} - \Pi^*_i A^*_{lk}\} - \Pi_r A_{rk} \Pi^*_s A^*_{sk} - \tfrac{3}{4}\kappa^2 C^* C + \tfrac{1}{2}\{\Pi_r A_{rk} \Pi^*_k C^* + \Pi^*_r A^*_{rk} \Pi_k C\} - \tfrac{3}{8}\Pi_l C \Pi^*_l C^*.$$

Performing the variations with respect to A^*_{ik} and C^* one obtains the equations

$$2\kappa^2 A_{ik} + 2\Pi^2 A_{ik} - 2\{\Pi_i \Pi_s A_{sk} + \Pi_k \Pi_s A_{si} - \tfrac{1}{2}\delta_{ik} \Pi_r \Pi_s A_{rs}\} + f_{ir} A_{rk} + f_{kr} A_{ri} + \tfrac{1}{2}\{\Pi_i \Pi_k + \Pi_k \Pi_i\} C - \tfrac{1}{4}\delta_{ik} \Pi^2 C = 0, \quad (15{\cdot}1)$$

$$-\tfrac{3}{2}\kappa^2 C - \tfrac{3}{4}\Pi^2 C + \Pi_r \Pi_s A_{rs} = 0. \quad (15{\cdot}2)$$

We shall now show that again 10 subsidiary conditions follow from these equations. We obtain three such conditions from (15·1) by putting $i \neq 4$, $k = 4$:

$$2\kappa^2 A_{\alpha 4} + 2\Pi^2_\beta A_{\alpha 4} - 2\Pi_\alpha \Pi_\beta A_{4\beta} - 2\Pi_\alpha \Pi_4 A_{44} - 2\Pi_4 \Pi_\beta A_{\beta\alpha} + f_{\alpha r} A_{r4} + f_{4\beta} A_{\beta\alpha} + \Pi_\alpha \Pi_4 C + f_{4\alpha} C = 0. \quad (16{\cdot}3)\ldots(16{\cdot}5)$$

$$(\alpha, \beta = 1, 2, 3; \quad r = 1, 2, 3, 4).$$

Adding the (4, 4)-component of (15·1) to (15·2) gives another:

$$2\kappa^2 A_{44} + 2\Pi^2_\beta A_{44} + 2\Pi_\alpha \Pi_\beta A_{\alpha\beta} - \Pi^2_\alpha C - \tfrac{3}{4}\kappa^2 C = 0. \quad (16{\cdot}6)$$

Applying the operator Π_k to (15·1) and using (15·2) gives four more:

$$2\kappa^2\Pi_i A_{ik}+3f_{il}\Pi_l A_{ik}+\frac{1}{i}\frac{\partial f_{il}}{\partial x_l}A_{ik}-3f_{ik}\Pi_s A_{si}-\frac{1}{i}\frac{\partial f_{rk}}{\partial x_i}A_{ri}$$
$$-\tfrac{3}{2}\kappa^2\Pi_k C+\tfrac{3}{2}f_{lk}\Pi_l C+\frac{1}{i}\frac{\partial f_{lk}}{\partial x_l}C=0, \quad (16\cdot 7)\ldots(16\cdot 10)$$
$$(k=1,2,3,4).$$

Applying Π_k to these equations once more gives

$$2\kappa^2\Pi_i\Pi_k A_{ik}-\tfrac{3}{2}\kappa^2\Pi^2C+\frac{3}{i}\frac{\partial f_{il}}{\partial x_k}\Pi_l A_{ik}-\frac{3}{i}\frac{\partial f_{ik}}{\partial x_k}\Pi_l A_{li}$$
$$-3f_{ik}f_{kl}A_{li}+\Pi_k\frac{1}{i}\frac{\partial f_{il}}{\partial x_l}A_{ik}-\Pi_k\frac{1}{i}\frac{\partial f_{rk}}{\partial x_l}A_{rl}+\frac{3}{2i}\frac{\partial f_{lk}}{\partial x_k}\Pi_l C$$
$$+\tfrac{3}{4}f_{lk}f_{kl}C+\Pi_k\frac{1}{i}\frac{\partial f_{lk}}{\partial x_l}C=0.$$

Using (15·2) we obtain from this the subsidiary condition

$$3\kappa^4C+\frac{2}{i}\frac{\partial f_{il}}{\partial x_k}\Pi_l A_{ik}-\frac{2}{i}\frac{\partial f_{il}}{\partial x_l}\Pi_k A_{ik}-2\frac{\partial^2 f_{il}}{\partial x_l\partial x_k}A_{ik}$$
$$-3f_{ik}f_{kl}A_{il}-\frac{1}{2i}\frac{\partial f_{lk}}{\partial x_k}\Pi_l C+\tfrac{3}{4}f_{kl}f_{lk}C=0. \quad (16\cdot 1)$$

The tenth condition is found by differentiating (16·1) with respect to the time. One then obtains an expression containing the second time-derivative of $C, A_{\alpha\beta}, A_{\alpha 4}$ and A_{44}, but these can be eliminated with the help of equations (15·1), (15·2) and the time-derivatives of (16·3)–(16·6). We shall not give these rather confusing calculations but content ourselves with the knowledge that the 10 subsidiary conditions exist.

The expression for the electric charge-current vector is obtained by forming the derivatives of L with respect to the four-potential. One finds

$$s_k=\frac{e}{hc}[i\{B_{[ki]l}A^*_{il}-B^*_{[ki]l}A_{il}\}+A^*_{lk}\Pi_r A_{rl}+A_{lk}\Pi^*_r A^*_{rl}$$
$$-\tfrac{1}{2}\{C^*\Pi_r A_{rk}+C\Pi^*_r A^*_{rk}+A^*_{kr}\Pi_r C+A_{kr}\Pi^*_r C^*\}$$
$$+\tfrac{3}{8}\{C^*\Pi_k C+C\Pi^*_k C^*\},$$

where $B_{[ik]l}\equiv i(\Pi_k A_{il}-\Pi_i A_{kl})$. For the force-free case, when Π_k becomes $-i\partial/\partial x_k$, this takes on the form given in (A). If one omits from L the term proportional to the field strengths one obtains an expression which is different from the above even in the force-free case, namely,

$$\bar{s}_k=\frac{e}{ihc}\left\{A^*_{il}\frac{\partial A_{il}}{\partial x_k}-A_{il}\frac{\partial A^*_{il}}{\partial x_k}\right\}.$$

This shows that the expression for the current in the limiting case of no forces is not unique, a point which incidentally arises already for the case of spin 1.

II. SPIN 3/2

4. Theory with no forces

In (A) it was shown that a force-free wave field corresponding to particles of spin 3/2 was described by spinors

$$a^{\dot\alpha}_{\beta\gamma} = a^{\dot\alpha}_{\gamma\beta}, \quad b^{\dot\alpha\dot\beta}_{\gamma} = b^{\dot\beta\dot\alpha}_{\gamma},$$

which go into one another by reflexion. They satisfy the equations

$$p^{\dot\beta\rho}a^{\dot\alpha}_{\rho\gamma} + p^{\dot\alpha\rho}a^{\dot\beta}_{\rho\gamma} = 2\kappa\, b^{\dot\alpha\dot\beta}_{\gamma}, \quad p_{\alpha\dot\rho}b^{\dot\rho\dot\gamma}_{\beta} + p_{\beta\dot\rho}b^{\dot\rho\dot\gamma}_{\alpha} = 2\kappa\, a^{\dot\gamma}_{\alpha\beta}, \tag{17·1}$$

together with the conditions

$$p_{\dot\alpha}{}^{\beta}a^{\dot\alpha}_{\beta\gamma} = 0, \quad p_{\dot\alpha}{}^{\gamma}b^{\dot\alpha\dot\beta}_{\gamma} = 0, \tag{17·2}$$

where $$p_{\dot\alpha\beta} = \sigma^{k}_{\dot\alpha\beta}\frac{1}{i}\frac{\partial}{\partial x_k}, \quad \text{with } \sigma^{k}_{\dot\alpha\beta} = (\boldsymbol{\sigma}^x, \boldsymbol{\sigma}^y, \boldsymbol{\sigma}^z, i\boldsymbol{I})_{\dot\alpha\beta}$$

(cf. also the explanations in the Appendix. We only wish to point out here that $\sigma_k^{\dot\alpha\beta}$ is the Hermitian conjugate of $-\sigma^k_{\dot\alpha\beta}$). The second order wave equation for $a^{\dot\gamma}_{\alpha\beta}$ and $b^{\dot\alpha\dot\beta}_{\gamma}$ follows from these equations.

The additional conditions (17·2) mean that no particles of spin 1/2 are to be present. Fields which contain particles of spin 1/2 as well as of spin 3/2 have no longer a definite form for the total charges in the c-number theory and so cannot be quantized in accordance with the exclusion principle. On the other hand the latter is physically necessary in order that the energy should be positive in the q-number theory (cf. (A)). Equations (17·1) and (17·2) can, as in the previous case, be derived from a variation principle if one introduces auxiliary variables c_α and $d^{\dot\alpha}$. A suitable choice of the constants in L again causes the quantities c_α, $d^{\dot\alpha}$, $p_{\dot\alpha}{}^{\beta}a^{\dot\alpha}_{\beta\gamma}$, $p_{\dot\alpha}{}^{\gamma}b^{\dot\alpha\dot\beta}_{\gamma}$ (which belong to the spin value $\frac{1}{2}$) to vanish as a consequence of the field equations in the force-free case. One has to choose for L the following:

$$\begin{aligned} L = \kappa\{a^{*\gamma}_{\dot\alpha\dot\beta}b^{\dot\alpha\dot\beta}_{\gamma} + b^{*\alpha\beta}_{\dot\gamma}a^{\dot\gamma}_{\alpha\beta}\} - \{a^{*\gamma}_{\dot\alpha\dot\beta}p^{\dot\beta\rho}a^{\dot\alpha}_{\gamma\rho} + b^{*\alpha\beta}_{\dot\gamma}p_{\alpha\dot\rho}b^{\dot\gamma\dot\rho}_{\beta}\} \\ + \{a^{*\gamma}_{\dot\alpha\dot\beta}p^{\dot\beta}{}_{\gamma}d^{\dot\alpha} + b^{*\alpha\beta}_{\dot\gamma}p^{\dot\gamma}{}_{\beta}c_\alpha + \text{conjugate}\} \\ + 3\{d^{*\alpha}p_{\alpha\dot\beta}d^{\dot\beta} + c^{*}{}_{\dot\alpha}p^{\dot\alpha\beta}c_\beta\} + 6\kappa\{d^{*\alpha}c_\alpha + d^{\dot\alpha}c^{*}{}_{\dot\alpha}\}. \end{aligned} \tag{18}$$

Variation gives the equations

$$\left.\begin{aligned} 2\kappa b^{\dot\alpha\dot\beta}_{\gamma} - p^{\dot\beta\rho} a^{\dot\alpha}_{\gamma\rho} - p^{\dot\alpha\rho} a^{\dot\beta}_{\gamma\rho} + p^{\dot\beta}{}_{\gamma} d^{\dot\alpha} + p_{\alpha\dot\beta} d^{\dot\beta} &= 0, \\ 2\kappa a^{\dot\gamma}_{\alpha\beta} - p_{\alpha\dot\rho} b^{\dot\rho\dot\gamma}_{\beta} - p_{\beta\dot\rho} b^{\dot\rho\dot\gamma}_{\alpha} + p^{\dot\gamma}{}_{\beta} c_{\alpha} + p^{\dot\gamma}{}_{\alpha} c_{\beta} &= 0. \end{aligned}\right\} \tag{19·1}$$

$$\left.\begin{aligned} -p^{\beta}{}_{\dot\gamma} a^{\dot\gamma}_{\alpha\beta} + 3p_{\alpha\dot\beta} d^{\dot\beta} + 6\kappa c_{\alpha} &= 0, \\ -p^{\gamma}{}_{\dot\beta} b^{\dot\alpha\dot\beta}_{\gamma} + 3p^{\dot\alpha\beta} c_{\beta} + 6\kappa d^{\dot\alpha} &= 0. \end{aligned}\right\} \tag{19·2}$$

One can now deduce from these equations that c_{α} and $d^{\dot\alpha}$ vanish, as well as that (17·2) is valid. To do this let us apply $p^{\gamma}{}_{\dot\beta}$ to the first of equations (19·1) and $p^{\alpha\dot\alpha}$ to the first of equations (19·2), remembering that

$$p_{\alpha\dot\beta} p^{\dot\beta\gamma} = \delta_{\alpha}{}^{\gamma}\Box, \quad p_{\alpha\dot\beta} p^{\dot\beta}{}_{\gamma} = \epsilon_{\alpha\gamma}\Box.$$

One finds

$$\left.\begin{aligned} 2\kappa p^{\gamma}{}_{\dot\beta} b^{\dot\alpha\dot\beta}_{\gamma} - p^{\dot\alpha\rho} p^{\gamma}{}_{\dot\beta} a^{\dot\beta}_{\gamma\rho} + 3\Box d^{\dot\alpha} \qquad\qquad &= 0, \\ - p^{\dot\alpha\rho} p^{\gamma}{}_{\dot\beta} a^{\dot\beta}_{\gamma\rho} + 3\Box d^{\dot\alpha} + 6\kappa p^{\dot\alpha\beta} c_{\beta} &= 0. \end{aligned}\right\} \tag{20}$$

By subtraction it follows that

$$2\kappa\{p^{\gamma}{}_{\dot\beta} b^{\dot\alpha\dot\beta}_{\gamma} - 3p^{\dot\alpha\dot\beta} c_{\beta}\} = 0.$$

Comparing this with the second of equations (19·2) we see that

$$d^{\dot\alpha} = 0.$$

Similarly one shows that the reflected quantity c_{α} vanishes. Equations (17·2) then follow from (19·2). Equations (19·1) take the form

$$\left.\begin{aligned} \kappa b^{\dot\alpha\dot\beta}_{\dot\gamma} &= p^{\dot\beta\rho} a^{\dot\alpha}_{\gamma\rho}, \\ \kappa a^{\dot\gamma}_{\alpha\beta} &= p_{\alpha\dot\rho} b^{\dot\rho\dot\gamma}_{\beta}, \end{aligned}\right\}$$

from which follows the wave equation for $a^{\dot\alpha}_{\beta\gamma}$ and $b^{\dot\alpha\dot\beta}_{\gamma}$.

5. Introduction of forces

One can now again introduce electromagnetic forces by replacing $p_{\alpha\dot\beta}$ by $\Pi_{\alpha\dot\beta}$, where $\Pi_{\alpha\dot\beta}$ is the spinor corresponding to the Π_k already defined. One then has the following equations:

$$\left.\begin{aligned} 2\kappa b^{\dot\alpha\dot\beta}_{\gamma} - \Pi^{\dot\beta\rho} a^{\dot\alpha}_{\gamma\rho} - \Pi^{\dot\alpha\rho} a^{\dot\beta}_{\gamma\rho} + \Pi^{\dot\beta}{}_{\gamma} d^{\dot\alpha} + \Pi^{\dot\alpha}{}_{\gamma} d^{\dot\beta} &= 0, \\ 2\kappa a^{\dot\gamma}_{\alpha\beta} - \Pi_{\alpha\dot\rho} b^{\dot\gamma\dot\rho}_{\beta} - \Pi_{\beta\dot\rho} b^{\dot\gamma\dot\rho}_{\alpha} + \Pi^{\dot\gamma}{}_{\beta} c_{\alpha} + \Pi^{\dot\gamma}{}_{\alpha} c_{\beta} &= 0, \end{aligned}\right\} \tag{21·1}$$

$$\left.\begin{aligned} -\Pi^{\beta}{}_{\dot\gamma} a^{\dot\gamma}_{\alpha\beta} + 3\Pi_{\alpha\dot\beta} d^{\dot\beta} + 6\kappa c_{\alpha} &= 0, \\ -\Pi^{\gamma}{}_{\dot\rho} b^{\dot\alpha\dot\beta}_{\gamma} + 3\Pi^{\dot\alpha\beta} c_{\beta} + 6\kappa d^{\dot\alpha} &= 0. \end{aligned}\right\} \tag{21·2}$$

Eight subsidiary conditions must follow from these equations, which in the force-free case must lead to the vanishing of c_{α}, $d^{\dot\alpha}$, and the validity of (17·2).

As the equations (21·1), (21·2) are of first order in the time derivatives the subsidiary conditions must not contain any time derivatives at all. Applying $\Pi^{\gamma}{}_{\dot\beta}$ to (21·1), $\Pi^{\alpha\dot\alpha}$ to (21·2) we find, analogously to the previous case,

$$2\kappa\Pi^{\gamma}{}_{\dot\beta}b^{\dot\beta\dot\alpha}_{\gamma}-\tfrac{1}{2}(\Pi^{\gamma}{}_{\dot\beta}\Pi^{\rho\dot\beta}+\Pi^{\rho}{}_{\dot\beta}\Pi^{\gamma\dot\beta})a^{\dot\alpha}_{\gamma\rho}-\Pi^{\gamma}{}_{\dot\beta}\Pi^{\dot\alpha\rho}a^{\dot\beta}_{\gamma\rho}+f_{\dot\beta}{}^{\dot\alpha}d^{\dot\beta}-3\Pi^{2}d^{\dot\alpha}=0,$$

$$-\Pi^{\dot\alpha\rho}\Pi^{\gamma}{}_{\dot\beta}a^{\dot\beta}_{\gamma\rho}+3f_{\dot\beta}{}^{\dot\alpha}d^{\dot\beta}-3\Pi^{2}d^{\dot\alpha}+6\kappa\Pi^{\rho\dot\alpha}c_{\rho}=0,$$

where we have substituted

$$\Pi_{\dot\alpha\beta}\Pi^{\dot\gamma\delta}-\Pi_{\dot\gamma\delta}\Pi^{\dot\alpha\beta}=\delta_{\dot\alpha}{}^{\dot\gamma}f_{\beta}{}^{\delta}+\delta_{\beta}{}^{\delta}f_{\dot\alpha}{}^{\dot\gamma},$$

$f^{\beta\delta}$, $f^{\dot\alpha\dot\gamma}$ being the symmetrical spinors corresponding to the electromagnetic field strengths.

Subtracting the two equations we obtain

$$2\kappa\Pi^{\gamma}{}_{\dot\beta}b^{\dot\beta\dot\alpha}_{\gamma}-2f^{\gamma\rho}a^{\dot\alpha}_{\gamma\rho}-2f_{\dot\beta}{}^{\dot\alpha}d^{\dot\beta}-6\kappa\Pi^{\rho\dot\alpha}c_{\rho}=0. \tag{22}$$

If we now compare this with the second of equations (21·2) we find

$$6\kappa d^{\dot\alpha}-\frac{1}{\kappa}f_{\dot\beta}{}^{\dot\alpha}d^{\dot\beta}=\frac{1}{\kappa}f^{\gamma\rho}a^{\dot\alpha}_{\gamma\rho}. \tag{23}$$

Similarly the reflected equation follows:

$$6\kappa c_{\alpha}-\frac{1}{\kappa}f_{\alpha}{}^{\beta}c_{\beta}=\frac{1}{\kappa}f_{\dot\gamma\dot\rho}b^{\dot\gamma\dot\rho}_{\alpha}. \tag{24}$$

This gives four subsidiary conditions. To be able to find four more we must separate time and space derivatives in (21·1), (21·2). For this we consider these equations in a co-ordinate system in which the time co-ordinate is fixed and therefore only require invariance under rotations of space. The spinor $s^{\dot\alpha}$ is then equivalent to s_{α}, and $\Pi^{\dot\alpha}{}_{\beta}$ is equivalent to $\Pi_{\alpha\beta}$, where

$$\Pi_{\alpha\beta}=\sum_{k=1}^{3}\sigma^{k}_{\alpha\beta}\Pi_{k}.$$

When written in this form the equations contain Π_4 explicitly as well. The first equations in (21·1), (21·2) become

$$\begin{aligned}2\kappa b_{\alpha\beta,\gamma}-\Pi_{\beta}{}^{\rho}a_{\alpha,\gamma\rho}-i\Pi_4 a_{\alpha,\gamma\beta}-\Pi_{\alpha}{}^{\rho}a_{\beta,\gamma\rho}-i\Pi_4 a_{\beta,\gamma\alpha}&\\+\Pi_{\beta\gamma}d_{\alpha}+i\epsilon_{\beta\gamma}\Pi_4 d_{\alpha}+\Pi_{\alpha\gamma}d_{\beta}+i\epsilon_{\alpha\gamma}\Pi_4 d_{\beta}&=0,\end{aligned} \tag{25}$$

$$-\Pi^{\beta\gamma}a_{\gamma,\alpha\beta}+i\Pi_4\epsilon^{\beta\gamma}a_{\gamma,\alpha\beta}+3\Pi_{\alpha}{}^{\beta}d_{\beta}+3i\Pi_4 d_{\alpha}+6\kappa c_{\alpha}=0. \tag{26}$$

We now multiply (25) by $\epsilon^{\beta\gamma}$ and obtain

$$\begin{aligned}2\kappa b_{\alpha\beta,}{}^{\beta}-\Pi_{\beta}{}^{\rho}a_{\alpha,\rho}{}^{\beta}-\Pi_{\alpha}{}^{\rho}a_{\beta,}{}^{\beta}{}_{\rho}-i\Pi_4 a_{\beta,}{}^{\beta}{}_{\rho}&\\+\Pi_{\beta}{}^{\beta}d_{\alpha}+3i\Pi_4 d_{\alpha}+\Pi_{\alpha}{}^{\beta}d_{\beta}&=0.\end{aligned} \tag{27}$$

Forming the difference of (27) and (26) we obtain

$$2\kappa b_{\alpha\beta,}{}^{\beta}+\Pi^{\beta\rho}a_{\alpha,\beta\rho}+\Pi_{\alpha\rho}a_{\beta,}{}^{\beta\rho}+\Pi^{\beta\gamma}a_{\gamma,\alpha\beta}$$
$$+\Pi_{\beta}{}^{\beta}d_{\alpha}-2\Pi_{\alpha}{}^{\beta}d_{\beta}-6\kappa c_{\alpha}=0. \tag{28}$$

Similarly for the reflected equation we find

$$2\kappa a_{\alpha\beta,}{}^{\beta}+\Pi^{\beta\rho}b_{\alpha,\beta\rho}+\Pi_{\alpha\rho}b_{\beta,}{}^{\beta\rho}+\Pi^{\beta\gamma}b_{\gamma,\alpha\beta}$$
$$+\Pi_{\beta}{}^{\beta}c_{\alpha}-2\Pi_{\alpha}{}^{\beta}c_{\beta}-6\kappa d_{\alpha}=0. \tag{29}$$

Equations (28) and (29) constitute 4 more subsidiary conditions. As a consequence of these 8 conditions the dimensionality of the states is unaltered by a field, and we can develop the theory in powers of the charge.

In conclusion we give the current-charge density vector

$$s_{\alpha\dot\beta}=a^{*\gamma}_{\dot\mu\dot\beta}a^{\dot\mu}_{\gamma\alpha}+b^{*}_{\dot\gamma,\alpha}{}^{\mu}b^{\dot\gamma}{}_{\dot\beta,\mu}$$
$$+\{a^{*}{}_{\dot\mu\dot\beta,\alpha}d^{\dot\mu}+b^{*}{}_{\dot\beta,\alpha\mu}c^{\mu}+\text{conjugate}\}-3\{d^{*}_{\alpha}d_{\dot\beta}+c^{*}_{\dot\beta}c_{\alpha}\}.$$

For the force-free case this reduces to the expression which was denoted by $s^{(0)}_{\alpha\dot\beta}$ in (A).

III. REST-MASS ZERO

6. Spin 2

One can set κ equal to zero in the formulae derived above for A_{ik} and C and so obtain a theory for zero rest-mass. The equations then run

$$-2\Box A_{ik}+2\left\{\frac{\partial^2 A_{sk}}{\partial x_s\,\partial x_i}+\frac{\partial^2 A_{si}}{\partial x_s\,\partial x_k}-\tfrac{1}{2}\delta_{ik}\frac{\partial^2 A_{rs}}{\partial x_r\,\partial x_s}\right\}-\frac{\partial^2 C}{\partial x_i\,\partial x_k}+\tfrac{1}{4}\delta_{ik}\Box C=0, \tag{30·1}$$

$$\tfrac{3}{4}\Box C-\frac{\partial^2 A_{rk}}{\partial x_r\,\partial x_k}=0. \tag{30·2}$$

It no longer follows from these equations that C and $\partial A_{ik}/\partial x_i$ vanish. Nevertheless, there are four identities which follow from them. For if we differentiate (30·1) with respect to x_i, (30·2) with respect to x_k, we obtain in either case

$$\tfrac{3}{4}\Box\frac{\partial C}{\partial x_k}-\frac{\partial}{\partial x_k}\frac{\partial^2 A_{rl}}{\partial x_r\,\partial x_l}=0, \quad (k=1,2,3,4). \tag{31}$$

Thus by subtracting the two we get identically zero for each value of k. As a result of these identities it is possible to construct quantities A^0_{ik}, C^0

from an arbitrary vector field f_i, which satisfy the field equations identically. Let us write

$$\left.\begin{aligned} A^0_{ik} &= \frac{\partial f_i}{\partial x_k}+\frac{\partial f_k}{\partial x_i}-\tfrac{1}{2}\delta_{ik}\frac{\partial f_l}{\partial x_l},\\ C^0 &= 2\frac{\partial f_l}{\partial x_l}. \end{aligned}\right\} \tag{32}$$

One can easily verify that the Lagrange function to which (11) reduces on setting κ equal to zero is altered by the "gauge transformation"

$$\left.\begin{aligned} A'_{ik} &= A_{ik}+A^0_{ik},\\ C' &= C+C^0, \end{aligned}\right\} \tag{33}$$

only to the extent of a complete differential.

The formulation in (A) is the same as the one here if one chooses the gauge in such a way that

$$\frac{\partial A_{ik}}{\partial x_i} = 0,$$

which is analogous to the Lorentz condition for the electromagnetic potentials. This condition restricts the gauge transformations to the group discussed in (A). The present scheme is identical with Einstein's "first approximation" of the gravitational equations.

Einstein (1916) considers the equations for the gravitational field in the cases when the deviations from a Euclidean metric are small quantities of the first order. We write

$$g_{ik} = \delta_{ik}+\gamma_{ik}; \quad \gamma_{ii} = \gamma.$$

Now let us write $$\gamma_{ik} = A_{ik}+\tfrac{1}{4}\delta_{ik}C; \quad \gamma = C.$$

We obtain the following differential equations for γ_{ik}:

$$\left.\begin{aligned} -\Box\gamma_{ik}-\frac{\partial^2\gamma}{\partial x_i\,\partial x_k}+\frac{\partial^2\gamma_{lk}}{\partial x_l\,\partial x_i}+\frac{\partial^2\gamma_{li}}{\partial x_l\,\partial x_k}+\tfrac{1}{2}\delta_{ik}\left\{\Box\gamma-\frac{\partial^2\gamma_{lr}}{\partial x_l\,\partial x_r}\right\} &= 0,\\ \Box\gamma-\frac{\partial^2\gamma_{lr}}{\partial x_l\,\partial x_r} &= 0. \end{aligned}\right\} \tag{34}$$

These equations are the same as those that Einstein gave for space containing no matter.

The gauge transformation (33) occurs in the gravitational theory as an infinitesimal co-ordinate transformation. When interactions with matter occur and it is no longer sufficient to restrict oneself to the linear terms the gauge group is altered. This keeps the dimensionality of the possible transformations unchanged; four functions of position always remain arbitrary. It is well known that the existence of an energy-momentum tensor is closely connected with the invariance of the gravitational theory under these transformations. Similarly, the gauge invariance of Maxwell's theory is connected with the conservation of charge.

7. Spin 3/2

Setting κ equal to zero in the equations for spin 3/2 gives

$$\left.\begin{aligned} -p^{\dot\beta\rho}a^{\dot\alpha}_{\gamma\rho}-p^{\dot\alpha\rho}a^{\dot\beta}_{\gamma\rho}+p^{\dot\beta}{}_{\gamma}d^{\dot\alpha}+p^{\dot\alpha}{}_{\gamma}d^{\dot\beta} &= 0,\\ -p_{\alpha\dot\rho}b^{\dot\rho\dot\gamma}_{\beta}-p_{\beta\dot\rho}b^{\dot\rho\dot\gamma}_{\alpha}+p^{\dot\gamma}{}_{\beta}c_{\alpha}+p^{\dot\gamma}{}_{\alpha}c_{\beta} &= 0;\end{aligned}\right\} \tag{35}$$

$$\left.\begin{aligned} -p^{\beta}{}_{\dot\gamma}a^{\dot\gamma}_{\alpha\beta}+3p_{\alpha\dot\beta}d^{\dot\beta} &= 0,\\ -p^{\gamma}{}_{\dot\beta}b^{\dot\alpha\dot\beta}_{\gamma}+3p^{\dot\alpha\beta}c_{\beta} &= 0.\end{aligned}\right\} \tag{36}$$

From these equations there follow four, or if we follow the procedure of Majorana (1937) and impose reality conditions on the field quantities, two identities. For, differentiating (35) with $p_{\dot\alpha}{}^{\gamma}$, (36) with $p^{\dot\beta\alpha}$ we find in both cases

$$-p^{\dot\beta\rho}p_{\dot\alpha}{}^{\gamma}a^{\dot\alpha}_{\gamma\rho}+3\Box d^{\dot\beta}=0 \quad (\dot\beta=1,\,2),$$

and similarly for the reflected equations. One can therefore find solutions with the help of spinor fields of first rank f_α, $g^{\dot\alpha}$, which satisfy the equations identically. Let us write

$$\left.\begin{aligned} a^{0\dot\gamma}_{\alpha\beta} &= p_{\alpha}{}^{\dot\gamma}f_{\beta}+p_{\beta}{}^{\dot\gamma}f_{\alpha}; \quad d^{0\dot\beta}=p^{\dot\beta\alpha}f_{\alpha};\\ b^{0\dot\alpha\dot\beta}_{\gamma} &= p_{\gamma}{}^{\dot\alpha}g^{\dot\beta}+p_{\gamma}{}^{\dot\beta}g^{\dot\alpha}; \quad c^{0}{}_{\beta}=p_{\beta\dot\alpha}g^{\dot\alpha}.\end{aligned}\right\}$$

The effect of the transformation $a'^{\dot\gamma}_{\alpha\beta}=a^{0\dot\gamma}_{\alpha\beta}+a^{\dot\gamma}_{\alpha\beta}, \ldots$ on the Lagrange function is again to add a complete differential.

Whereas the theory for the spin value 2 has an important generalization for force fields, namely the gravitational theory, we here have no such connexion with a known theory. To get a generalization of the theory with interactions, one would first of all have to find a physical interpretation of the gauge group, and of the conservation theorem connected with this group.

8. General case of arbitrary spin

To set up a theory with forces for particles of arbitrary spin one again first looks for a variation principle from which the equations of (A) can be derived. The forces can then be introduced by suitable modifications of the

Lagrange function. For instance, the effect of an electromagnetic field can be described by replacing $-i\partial/\partial x_k$ by Π_k. In generalizing the method which we have already used for the spin values 2 and 3/2 we must again introduce auxiliary fields, which, in the force-free case, vanish as a consequence of the field equations. To illustrate the method it will be sufficient to discuss the case of integral spin.

We start from a tensor $A^f_{ik\ldots l}$, of rank f, symmetrical in all its indices and whose trace $A^f_{ii\ldots l}$ vanishes. Let us further introduce auxiliary fields of rank $f-2, \ldots, s, \ldots, 1, 0$ which are likewise all symmetrical and of zero trace; there may be several fields with the same rank, which we shall distinguish by an index t. The general field occurring in the Lagrange function we shall thus denote by $A^{st}_{ik\ldots l}$, $i, k, \ldots, l$ being tensor indices, with respect to which the field is symmetrical, and t distinguishes the different fields of rank s. The index s takes the values $0, 1, \ldots, f-2, f$. In general with a field A^s of rank s one can associate $s+1$ kinds of particles, i.e. those of spin s, $s-1, \ldots, 1, 0$. The Lagrangian must be so constructed that in the end only those of spin f occur. Or we can say that as a consequence of the field equations all the fields corresponding to particles of spin $f-1, \ldots, 1, 0$ vanish. For L we choose the following:

$$L = \sum_s \sum_t \left[\kappa^2 A^{st}_{ik\ldots} A^{st}_{ik\ldots} + a_1^{st} \frac{\partial A^{st}_{ik\ldots}}{\partial x_l} \frac{\partial A^{st}_{ik}}{\partial x_l} + a_2^{st} \frac{\partial A^{s+1,t}_{ik\ldots}}{\partial x_i} \frac{\partial A^{s+1,t}_{lk\ldots}}{\partial x_l} \right.$$
$$\left. + \sum_r \left\{ a_3^{s,tr} \frac{\partial A^{st}_{ik\ldots}}{\partial x_l} A^{s+1,r}_{il\ldots} + a_4^{s,tr} \frac{\partial A^{st}_{il\ldots}}{\partial x_k} \frac{\partial A^{s+2,r}_{mikl\ldots}}{\partial x_m} \right\} \right], \quad (37)$$

where $a_1^{st}, a_2^{st}, a_3^{st,r}, a_4^{st,r}$ are constants which must be so chosen that particles of spin s do not occur. In order to get the requisite number of constants we introduce auxiliary fields whose number is given by the following table:

s	Number of fields of rank s	Number of particles of spin s to be removed	Corresponding number of constants at our disposal
$f-1$	0	1	1
$f-2$	1	2	2
$f-3$	1	3	3
$f-4$	1	4	4
$f-5$	2	6	7
$f-6$	3	9	14
...	...	...	...
$f-n$	$n-3$	$\frac{1}{2}(n-2)(n-3)+3$	$2(n-3)^2-n+2$

One sees from the table that in this way there are more constants than are necessary. One can therefore construct the Lagrange function in many ways such that the same force-free theory follows as in (A). We have not yet been able to find any simple way of avoiding this ambiguity. In particular, it is impossible to manage with only one field of each rank, for one would then obtain n particles of spin $f-n$ and for $n \geqslant 5$ only 4 constants to remove them.

A completely analogous procedure also works for half-integral spin. The only difference lies in the fact that the Lagrangian contains derivatives of first order only, leading to a slight difference in the number of constants. It is to be remarked that already for 5/2 one must introduce *two* fields corresponding to spin 1/2. As the method is otherwise the same as for integral spin there is no need to go into the details.

Appendix

(1) *Forms of Lagrangian leading to singular solutions*

As we have already stressed, if one wants to modify the Lagrange function of the field for particles of spin $\geqslant 1$ in a manner corresponding to the interaction with other fields, one must take care that the number of restrictive conditions is not diminished. For then the switching on of the forces would create new particles whose corresponding particular solutions of the equations become singular when the field is switched off again.

As an example of this we shall give an inadmissible form of interaction of particles of spin 1 with a scalar field ψ. Let us write

$$L = \kappa^2 A_i^2 + \left(\frac{\partial A_i}{\partial x_k}\right)^2 - \left(\frac{\partial A_i}{\partial x_i}\right)^2 + \psi\left(\frac{\partial A_i}{\partial x_i}\right)^2. \tag{38}$$

The field equations are

$$\Box A_k - \frac{\partial^2 A_i}{\partial x_i \partial x_k} = \kappa^2 A_k - \frac{\partial}{\partial x_k}\left(\psi \frac{\partial A_i}{\partial x_i}\right). \tag{39}$$

From these follows
$$\frac{\partial A_i}{\partial x_i} = \Box\left(\psi \frac{\partial A_i}{\partial x_i}\right).$$

When $\psi \equiv 0$ this is a subsidiary condition for A_i which is, however, removed by the interaction.

To study the character of the new solutions arising from the interaction with ψ let us make the assumption that A_i and ψ depend only on the time. Then we have

$$\frac{\partial^2 A_k}{\partial x_4{}^2} - \delta_{k4} \frac{\partial^2 A_4}{\partial x_4{}^2} = \kappa^2 A_k - \delta_{k4} \frac{\partial}{\partial x_4}\left(\psi \frac{\partial A_4}{\partial x_4}\right).$$

The equation for A_4 is of interest since it contains ψ. It runs

$$\kappa^2 A_4 = \frac{\partial}{\partial x_4}\left(\psi \frac{\partial A_4}{\partial x_4}\right). \tag{40}$$

Let us suppose that ψ changes so slowly that $\partial\psi/\partial x_4$ can be neglected. We have then approximately

$$\frac{\kappa^2}{\psi} A_4 = \frac{\partial^2 A_4}{\partial {x_4}^2}.$$

The waves A_4 therefore correspond to a mass $\kappa/\sqrt{\psi}$. This mass is imaginary for negative ψ; for $\psi = 0$ it becomes infinite. One therefore obtains strongly singular solutions in the limit $\psi \to 0$. Equation (40) can be solved explicitly if one takes for ψ the form

$$\psi = \pm e^{-\alpha t}, \quad \alpha > 0.$$

One finds
$$A_4 = \mathfrak{I} \frac{\text{const.}}{\sqrt{\psi}} Z_1(2\kappa/\alpha\sqrt{\psi}),$$

where $\mathfrak{I}$ denotes the imaginary part, and Z_1 is a Bessel function of first order. A_4 diverges exponentially as ψ approaches zero from negative values.

(2) *Rules for spinor calculus*

In the following we shall collect a few definitions and rules of spinor calculus which have been used in the previous pages.

(1) Spinor indices are raised and lowered according to the following rule:

$$v_1 = -v^2; \quad v_2 = v^1. \tag{1}$$

The scalar product of two spinors is accordingly

$$v_\alpha u^\alpha = -v^\alpha u_\alpha = v_1 u_2 - v_2 u_1. \tag{2}$$

This can be expressed in terms of the invariant spinor

$$\epsilon_{12} = -\epsilon_{21} = \epsilon^{12} = -\epsilon^{21} = 1, \quad \epsilon_{11} = \epsilon_{22} = 0.$$

$$v^\alpha = \epsilon^{\alpha\beta} v_\beta, \quad v_\alpha = v^\beta \epsilon_{\beta\alpha},$$

$$v_\alpha u^\alpha = \epsilon^{\alpha\beta} v_\alpha u_\beta = -\epsilon^{\alpha\beta} u_\alpha v_\beta.$$

Thus to raise a suffix one applies $\epsilon^{\alpha\beta}$ on the left, to lower, on the right.

The transition from spinors to four-dimensional tensors is done by means of the matrices $\sigma^k_{\dot\alpha\beta}$, where k runs from 1 to 4, α and β from 1 to 2. They are defined by

$$\sigma^1_{\dot\alpha\beta} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma^2_{\dot\alpha\beta} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma^3_{\dot\alpha\beta} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \sigma^4_{\dot\alpha\beta} = \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix}. \tag{3}$$

(This differs from the notation of v. d. Waerden (1932), who uses $\sigma^4 = I$. The above is more convenient here as we are using the imaginary co-ordinate $x_4 = ict$.) From (1) it follows that

$$\sigma^k_{1\dot{2}} = -\sigma^{k,\dot{1}2}, \quad \sigma^k_{2\dot{1}} = -\sigma^{k,\dot{2}1},$$
$$\sigma^k_{1\dot{1}} = \sigma^{k,\dot{2}2}, \qquad \sigma^k_{2\dot{2}} = \sigma^{k,\dot{1}1}.$$

As the trace of σ^k vanishes for $k = 1, 2, 3$ we have the following rule:

$$\sigma^{k,\dot{\alpha}\beta} = -\sigma^k_{\dot{\alpha}\beta} \quad \text{for } k = 1, 2, 3,$$
$$\sigma^{4,\dot{\alpha}\beta} = \sigma^4_{\dot{\alpha}\beta}.$$

The σ^k satisfy the commutation rules

$$\left.\begin{aligned} \sigma^k_{\dot{\alpha}\beta}\sigma^{l,\beta\dot{\gamma}} + \sigma^l_{\dot{\alpha}\beta}\sigma^{k,\beta\dot{\gamma}} &= -2\delta_{kl}\delta_{\dot{\alpha}}{}^{\dot{\beta}}, \\ \sigma^k_{\dot{\alpha}\beta}\sigma^{l,\gamma\dot{\alpha}} + \sigma^l_{\dot{\alpha}\beta}\sigma^{k,\gamma\dot{\alpha}} &= -2\delta_{kl}\delta_{\beta}{}^{\gamma}. \end{aligned}\right\} \tag{4}$$

One puts a four-vector a_k in correspondence with a spinor $a_{\dot{\alpha}\beta}$ with the help of the $\sigma^k_{\dot{\alpha}\beta}$ as follows:

$$a_{\dot{\alpha}\beta} = a_k \sigma^k_{\dot{\alpha}\beta}. \tag{5}$$

Conversely we have

$$a_{\dot{\alpha}\beta}\sigma^{k,\dot{\alpha}\beta} = -2a_k. \tag{6}$$

The four-vector a_k can also be an operator, for instance $-i\partial/\partial x_k$. If the components of a_k commute with one another,

$$a_i a_k - a_k a_i = 0,$$

then it follows from (4) that

$$\left.\begin{aligned} a_{\dot{\alpha}\beta}a^{\beta\dot{\gamma}} &= -a^2\delta_{\dot{\alpha}}{}^{\dot{\gamma}}, \\ a_{\dot{\alpha}\beta}a^{\gamma\dot{\alpha}} &= -a^2\delta_{\beta}{}^{\gamma}, \end{aligned}\right\} \tag{7}$$

where

$$a^2 = \sum_{k=1}^{4} a_k^2 = -\tfrac{1}{2}a_{\dot{\alpha}\beta}a^{\beta\dot{\alpha}}.$$

In particular for the spinor

$$p_{\dot{\alpha}\beta} = \frac{1}{i}\frac{\partial}{\partial x_k}\sigma^k_{\dot{\alpha}\beta}, \tag{8}$$

it follows from (7) that

$$p_{\dot{\alpha}\beta}p^{\beta\dot{\gamma}} = \delta_{\dot{\alpha}}{}^{\dot{\gamma}}\Box, \tag{9}$$

where

$$\Box \equiv \sum_{k=1}^{4}\frac{\partial^2}{\partial x_k^2}.$$

Again, if

$$\Pi_k = \frac{1}{i}\frac{\partial}{\partial x_k} - \frac{e}{hc}\phi_k, \tag{10}$$

then we have

$$\Pi_k\Pi_l - \Pi_l\Pi_k = \frac{ie}{hc}\left(\frac{\partial\phi_l}{\partial x_k} - \frac{\partial\phi_k}{\partial x_l}\right) \equiv f_{kl}. \tag{11}$$

As the components of Π_k do not commute with one another we have, from (4)

$$\begin{aligned}\Pi_{\dot{\alpha}\beta}\Pi^{\beta\dot{\gamma}} &= \sigma^k_{\dot{\alpha}\beta}\sigma^{l,\,\beta\dot{\gamma}}\Pi_k\Pi_l \\ &= \tfrac{1}{2}\sigma^k_{\dot{\alpha}\beta}\sigma^{l,\,\beta\dot{\gamma}}\{\Pi_k\Pi_l+\Pi_l\Pi_k+f_{kl}\} \\ &= -\Pi^2\delta_{\dot{\alpha}}{}^{\dot{\gamma}}+f_{\dot{\alpha}}{}^{\dot{\gamma}}, \end{aligned} \tag{12}$$

where

$$\begin{aligned}\sigma^k_{\dot{\alpha}\beta}\sigma^{l,\,\dot{\gamma}\delta}f_{kl} &= \delta_\beta{}^\delta f_{\dot{\alpha}}{}^{\dot{\gamma}}+\delta_{\dot{\alpha}}{}^{\dot{\gamma}}f_\beta{}^\delta \\ &= \Pi_{\dot{\alpha}\beta}\Pi^{\dot{\gamma}\delta}-\Pi^{\dot{\gamma}\delta}\Pi_{\dot{\alpha}\beta}\end{aligned} \tag{13}$$

is the spinor corresponding to the field strengths (cf. Uhlenbeck and Laporte 1931), with the property that

$$f_\alpha{}^\alpha = f_{\dot{\alpha}}{}^{\dot{\alpha}} = 0.$$

Besides the commutation rules (4) the $\sigma^k_{\dot{\alpha}\beta}$ also satisfy the relations

$$\sum_{k=1}^{4}\sigma^k_{\dot{\alpha}\beta}\sigma^k_{\dot{\gamma}\delta} = 2(\delta_{\dot{\alpha}\delta}\delta_{\dot{\gamma}\beta}-\delta_{\dot{\alpha}\beta}\delta_{\dot{\gamma}\delta}) = 2\epsilon_{\dot{\alpha}\dot{\gamma}}\epsilon_{\delta\beta}. \tag{14}$$

Let us now consider the tensor b_{kl} of second rank corresponding to the symmetrical spinor

$$b_{\dot{\alpha}\dot{\beta},\,\gamma\delta} = b_{\dot{\beta}\dot{\alpha},\,\gamma\delta} = b_{\dot{\alpha}\dot{\beta},\,\delta\gamma}.$$

It follows at once from (14) that its trace vanishes,

$$b_{kk} = 0.$$

Summary

The force-free theory of particles with arbitrary spin values already published by one of the authors is generalized to the relativistic wave equations of such particles in an electromagnetic field, with a preliminary restriction to the c-number theory. The spin values 3/2 and 2 are treated in detail, and for the general case it is merely proved that consistent wave equations exist. The consistency of the system of field equations is attained by deriving them from a Lagrange function containing suitable additional terms which depend on new auxiliary quantities. All the differential equations of the field are derived by variation of the action integral and the vanishing of the auxiliary quantities in the absence of an external field is made to follow as a consequence of them.

In the special case of zero rest-mass there exist identities between the equations, which are now invariant under a group of transformations which is the generalization of the group of gauge transformations in Maxwell's

theory. In the particular case of spin 2, rest-mass zero, the equations agree in the force-free case with Einstein's equations for gravitational waves in general relativity in first approximation; the corresponding group of transformations arises from the infinitesimal co-ordinate transformations.

References

Dirac 1936 *Proc. Roy. Soc.* A, **155**, 447.
Einstein 1916 *S.B. preuss. Akad. Wiss.* (*Math. Phys.*), p. 688.
— 1918 *S.B. preuss. Akad. Wiss.* (*Math. Phys.*), p. 154.
Fierz 1939 *Helv. Phys. Acta*, **12**, 3 (referred to as (A)).
Majorana 1937 *Nuovo Cim.* **14**, 171.
Uhlenbeck and Laporte 1931 *Phys. Rev.* **37**, 1380.
v. d. Waerden 1932 *Die gruppentheoretische Methode in der Quantenmechanik*, p. 83. Berlin: Springer.

Defect lattices in some ternary alloys

By H. Lipson, D.Sc.,

Crystallographic Department, Cavendish Laboratory, Cambridge

and A. Taylor, Ph.D.,

Chemistry Department, King's College, Newcastle-upon-Tyne

(*Communicated by A. J. Bradley, F.R.S.—Received* 9 *June* 1939)

In the work on ternary alloys with which the authors have been associated (Bradley, Goldschmidt, Lipson and Taylor 1937), the most outstanding feature has been the large range of composition over which the β (body-centred cubic) structure exists. This was not unexpected in the system Fe-Ni-Al (Bradley and Taylor 1938), where FeAl and NiAl both have this structure and thus may be expected to be isomorphous; but CuAl has quite a different structure, and it was therefore surprising to find the β phase-fields extending towards this composition in the systems Cu-Ni-Al (Bradley and Lipson 1938) and Fe-Cu-Al (Bradley and Goldschmidt 1939).

These extensive phase-fields provide an opportunity for a more detailed examination of the phenomenon of the occurrence of defect lattices which takes place in the Ni-Al system (Bradley and Taylor 1937). In the β

Discreteness and Divergences

Chapter 24
The Emergence of Quantum Geometry
Dean Rickles

A central theme of the previous selection of papers was the existence of a variety of divergences, especially the problems generated by the self-energy of elementary particles.[1] A common thread that emerges, persisting to this day, is that gravitation might be able to get some kind of foothold in the theory of elementary particles if it could be invoked to provide a physical (non-ad hoc) cutoff at wavelengths that do not conflict with current observations of experimental particle physics data.[2] Hence, Einstein's old question about the role of gravity in elementary particle physics (mentioned in the introduction to part one of this volume) comes alive once again, though not in the way he envisaged it: gravity is rather "put to service" in elementary particle physics as a kind of external resource. However, the notion of a *gravitationally* induced cutoff took some time to form, and had to wait for a more thorough understanding of the non-linear aspects of the theory, which had remained largely buried under the more tractable linear approximation.[3] But this work can be seen as emerging from a well-worn path involving the usage of other cutoffs (minimum lengths) to tame the infinite behaviour of field theories—if gravity suggests a minimum length, then previous work on such ideas can be transferred across from context to context. The papers presented in this part, then, as with many others in this book, were not initially written with quantum gravity in mind; this was a connection only made later once the problem of quantum gravity itself had undergone several changes—not least the idea that quantum gravity involves quantum spacetime.

There were many early suggestions that gravity might act as some kind of divergence "regulator". The divergences in question were those of QED, and meson theories, which were still, pre-WWII, a somewhat mathematically murky territory (quantum field theory still is, of course). The problem concerned the transitions between quantum states, during which time (a very short time, determined by the uncertainty relations) energy conservation is violated. The great hope for introducing gravitation into elementary particle physics was that it would terminate the wavelengths before they have a chance to reach the problematic high-energy (ultraviolet) wavelengths. Landau (1955) appears to be the first to have suggested this idea based on his more general desire to achieve something *beyond* quantum field theory, since Landau had a deep distrust of quantum field theory. More specifically, Landau's field theoretic investigations revealed that even given renormalization, the short-distance behaviour continued to generate infinities ("Landau ghosts") on account of the approach to the infinite bare masses and couplings (past the "screening"

[1] A good philosophical treatment of the reactions to these divergences is Rueger (1992).

[2] The alternative, to be found in superstring theory, is to postulate a finite size for the elementary entities (in the earlier period, this would have been, e.g., the electron radius).

[3] What the non-linear analysis eventually revealed was that gravity's divergences were more complex than those of electrodynamics, since they involved a shifting of the light cone so that light cone singularities would be "smoothed out" by fluctuations in a quantum theory of gravity (a nice result of the quantum geometry inherent in quantization of the metric): Klein 1956; Landau 1955; Deser 1957. We return to this briefly below.

effect that Landau had himself discovered).[4] Pauli also makes several comments to this effect,[5] including the following remarks in a letter addressed to Abrikosov, Khalatnikov, and Pomeranchuk (in which he assigns priority of the idea to Landau):

> I was very interested in *Landau's* remarks on the possibility of a connection of the cut-off moment of quantum electrodynamics with *gravitational* interaction (his article "on quantum theory of fields" in the Bohr-festival volume). It appeals [sic] to me, that the situation regarding divergencies would be fundamentally changed, as soon as the light-cone itself is not any longer a c-number equation. Then every given direction in space-time would have some "probability to be on the light-cone", which would be different from zero for a small but finite domain of directions. I doubt, however, that the *conventional* quantization of the $g_{\mu\nu}$-field is consistent under this circumstances. (Zürich, 15 August 1955; in Meyenn (2001, 329))

Given his obvious expertise in both general relativity and quantum field theory, one wonders why he didn't do more work in the area of quantum gravity. Though as he suggests, it is clear that novel (i.e. unconventional) approaches are probably required (a far cry from his optimistic remarks in the 1929 paper with Heisenberg in which he claims that the general relativistic case would be much the same as electromagnetism). He was, nonetheless, certainly preoccupied with general relativity towards the end of his life—perhaps this "later-life" preoccupation with the unification of quantum and gravity (also shared by Eddington and Schrödinger) contributed in some small way to the poor reputation of the field at this stage.

Developing his earlier remarks to Abrikosov et al. (in the discussion after Klein's talk), Pauli mentions Landau's argument that for large cutoff momentum P, the gravitational coupling between a pair of electrons is of the same magnitude as the Coulomb forces. He notes that Landau's relation $GP^2 \sim 1$ is the same as Klein's $P \sim \hbar/l_0$ (where $l_0 = \sqrt{G\hbar c}$). He writes:

> [T]he connection [...] of the mathematical limitation of quantum electrodynamics with gravitation, pointed out by LANDAU and KLEIN, seems to me to hint at the indeterminacy in space-time of the light-cone, which is governed by probability laws in a quantized field theory, invariant with respect to the wider group of general relativity. It is possible that this new situation so different from quantized theories, invariant with respect to the LORENTZ group only, may help to overcome the divergence difficulties which are so intimately connected with a c-number equation for the light-cone in the latter theories (Pauli's comments after Klein's talk, in Mercier and Kervaire 1956, 69).

Pauli's thoughts were borne out in one way (the light cone structure is affected as he suspects); however, it leaves a challenge behind in dealing with new divergences. As Bryce DeWitt pointed out, in Louis Witten's important collection, *Gravitation: An Introduction to Current Research*, from 1962 (less than seven years after Pauli's remarks),

[4] Later work would reveal that such Landau ghosts could be dealt with in the context of renormalization group theory, but at the time it appeared as though quantum field theory was suffering from an incurable illness—for more on this, see Brown (1993, 21).

[5] See also Pauli's letter to Källén (dated 24th April, 1955: pp. 207–208); Peierls' letter to Pauli (dated 9th May, 1955: pp. 228–229); Pauli's letter to van Hove (dated 11th May, 1955: 230–231); Heisenberg's letter to Pauli (dated 11th, May 1955: pp. 234–235).

> [I]t must constantly be borne in mind that the "bad" divergences of quantum gravidynamics are of an essentially different kind from those of other field theories. They are direct consequences of the fact that the light cone itself gets shifted by the non-linearities of the theory. But the light-cone shift is precisely what gives the theory its unique interest, and a special effort should be made to separate the divergences which it generates from other diverg'ences. (DeWitt 1962, 374).

In the United States, the first PhD thesis to be written on quantum gravity was that of Bryce DeWitt, under the supervision of Julian Schwinger (at Harvard University, completed in 1949). DeWitt sought to revisit Rosenfeld's work on the computation of gravitational self-energies (cf. Deser 1957). DeWitt would also revisit this idea of Landau's that gravity might act as a natural regulator (DeWitt 1964). Though Landau didn't explicitly mention the Planck scale (he placed the location of the cutoff much higher), Pauli clearly appeared to think that Landau had quantum gravitational effects in mind (or that he *ought* to have). It is clear that if there is a "fundamental length," below which ordinary quantum field theoretic processes cannot operate, then one has what Landau sought. DeWitt was able to confirm that (at lowest order of perturbation) when gravity is included, the self-energies of charged particles (and the gravitons themselves) remain finite (though often very large). Here again, as in earlier parts, we see a link between minimum length scales and the notion of *limits* and *domains of applicability* of theory and concepts. The question of whether there is a physical cutoff naturally has theoretical links with programmes concerning the existence of a fundamental length, and discrete space(time) in general.[6]

More indirect, however, was Peter Bergmann's method of utilising the fact that the gravitational field equations determined particle trajectories free of any notions of divergences. He believed this would follow from the analysis of Einstein, Hoffmann and Infeld, according to which the assumption of geodesy for a free particle's motion was redundant, since it already could be seen to follow (by a method of successive approximation) from the field equations alone.[7]

Developing the cutoff idea, and the idea that there might be a minimal (fundamental) length, leads one quite naturally into the idea that space and time might not be continuous,

[6] As DeWitt puts it: "The dimension $10^{-32}cm$ constitutes a fundamental limit on the smallness of allowable measurement domains. Below this limit it is impossible to interpret the results of measurement in terms of properties or states characterising individual systems under observation" (DeWitt 1962, 373).

[7] I might also note here that ultimately string theory emerged from the divergences problems facing quantum field theories of fields other than the electromagnetic field (particularly the strong interaction). In particular, since the perturbative approach breaks down when the coupling constant determining the strength of an interaction is large (as in strong interaction physics), alternative approaches were sought in the late 1950s and throughout the 1960s. One of the more popular of these approaches combined Heisenberg's S-matrix theory with dispersion theory. The S-matrix is a tool to encode all possible collision processes. Heisenberg suggested that one take this to embody what was relevant about the physics of collision processes. In particular, all that was observable were the inputs and outputs of collision processes, observed when the particles are far enough apart in spacetime to be non-interacting, or free. This black box approach to physics was very much inspired by the Copenhagen philosophy. The dispersion relation approach to physics tried to construct physical theories on the basis of a few central physical axioms, such as unitarity (conservation of probabilities), Lorentz invariance, and causality (effects can't precede causes). These two approaches were combined, by Geoff Chew amongst others, so that the focus was on the analytic properties of the S-matrix. One model for the S-matrix, incorporating some other principles thought to be involved in strong interaction physics, was the Veneziano model. This used the Euler beta function to encode the various desirable properties of the S-matrix. The model was found to be generated by a dynamical theory of strings. (See Cushing (1990) for a detailed historico-philosophical account of the early development of string theory, or Rickles (2014) for a more recent account.)

but better modelled instead by a discrete lattice or similar structure. In the early days the cutoff was implemented in the kinematical structure, rather than having it emerge dynamically—whether the cutoff (discreteness) is fundamental or not is a different issue. This was suggested by several people. In a paper from 1930 Ambarzumian and Iwanenko (Chapter 25) argued for the introduction of a spatial lattice structure for physical space as a way of eliminating the infinite divergences from the self-energy of the electron. The basic idea was that the existence of a minimal length would imply a maximal frequency (p. 416). Alfred Schild (Chapter 30) investigated the properties of such a discrete lattice in order to see if it would break essential symmetries. In particular, he was responding to the objection that discrete theories would violate Lorentz invariance, which could manifest experimentally resulting in inconsistencies with known results.[8] He wasn't able to devise a model to preserve all such symmetries, but enough to provide a plausible candidate for a background for a physical theory. Here again we find constraints operating on the various approaches to provide some sort of mechanism for the rejection and selection of theories or approaches—in this case the Lorentz symmetry of the classical theory.

Another discrete approach, of David van Dantzig (1938; 1956), was motivated by a combination of general covariance (as expressed in Einstein's "point-coincidence" argument) and the definition of observability in such a theory. He argued that in a generally covariant theory the observable things will be coincidences: events (not shuffled by diffeomorphisms). Van Dantzig argues that in order to not introduce unmeasurable structure into the interpretation or formulation of one's theory, one should dispense with the existence of a four-dimensional continuum, in favour of a discrete manifold of events. Peter Bergmann describes one such approach as one of "constructing "spaces" that have certain topological properties similar to those of point spaces in the large but do not possess "points" as elementary constituents" (Bergmann, following a talk of Wigner's: Wigner (1956, 226)). The general approach lives on in several of the current approaches, including causal set

[8]This same objection to discrete models surfaces again in present-day discussions of discrete space in quantum gravity (a fairly generic prediction of several approaches), especially in the context of loop quantum gravity which directly predicts (at least at the kinematic level) geometrical operators with a discrete spectrum. Given that there is supposed to be a fundamental length (namely the Planck length, and corresponding fundamental times and masses) in these approaches, it makes sense to ask if observers in relative motion will agree on this length: why no Lorentz-FitzGerald contraction for boosted observers, rendering the notion of a minimum length incoherent? Why is a length measurement for the minimum length case not subject to the usual frame dependence? According to Carlo Rovelli (one of the primary architects of loop quantum gravity that itself appears to face the problem) and Simone Speziale, quantum mechanics is the key to avoiding this "discreteness/invariance" conflict: "the minimal length (more precisely, minimal area) does not appear as a fixed property of geometry, but rather as the minimal (nonzero) eigenvalue of a quantum observable [so that the] boosted observer can see the same observable spectrum, with the same minimal area. What changes continuously in the boost transformation is not the value of the minimal length: it is the probability distribution of seeing one or the other of the discrete eigenvalues of the area" (Rovelli and Speziale 2003, 064019). They elaborate as follows, linking directly with issues of quantum spacetime: "The geometry of space comes from a quantum field, the quantum gravitational field. Therefore the observable properties of the geometry, such as, in particular, a length, or an area, are observable properties of a *quantum physical system*. A measurement of a length is therefore a measurement in the quantum mechanical sense. Generically, quantum theory does not predict an observable value: it predicts a probability distribution of possible observable values. Given a surface moving in spacetime, the two measurements of its area performed by two observers O and O' boosted with respect to one another are two entirely distinct quantum measurements. Correspondingly, in the theory there are two distinct operators A and A', associated to these two measurements. Now, our main point is the technical observation that A and A' do not commute:

$[A, A'] = 0$. This is because A and A' depend on the gravitational field on two distinct 2d surfaces in spacetime […] and a field operator does not commute with itself at different times". Hagar (2014, §8.4.4) contains a useful, detailed discussion of this problem.

theory and dynamical triangulations—though the conceptual basis (especially *observability through invariance* is absent from the latter case). Bergmann's comments also draw attention to the "emergence" of continuous spacetime from a discrete structure (a problem at the root of causal set theory, though one in which progress has been made: see, e.g., Major, Rideout, and Surya (2007)).

There was nothing corresponding to paradigms in the early work. Nobody pursued a single programme for long enough—though Bergmann's initial canonical quantization approach spawned a genuine research programme (along with a family of characteristic questions, having to do with "true observables" and the like) that has persisted. However, in the present day we do have a situation of what seem to be competing, coexisting paradigms (with elements of this sourcebook's papers as ingredients). We can find the seeds of this landscape in the emergence of various "schools of research", each tackling the problem of quantum gravity in a unique way. Often these schools themselves had seeds in the distinct *tools* that the researchers brought from their training, as physicists and mathematicians (recall that before the 1960s, it was rare to find general relativity taught outside of mathematics departments).

The idea at the focus of the papers in this part, that a discretization of space might go some way towards resolving the problems of short-distance physics, is of course rather natural and almost obvious. However, the initial developments were not linked to gravitational physics, although many of the results originally couched in non-gravitational work were carried over into that area. It was eventually realised, for example, that gravitation itself might be able to provide a *physical foundation* for discrete space and that given the dual nature of the metric field, quantum gravity should lead one to expect a discrete spacetime. Given this, the various results pursued independently of the quantum gravity problem (violation of Lorentz invariance and so on), become directly relevant.

There are three motivations underlying the notion of discrete space(time) in the early work:[9]

1. An *ad hoc* discretization using a lattice structure—often used as an approximation, for which the continuum limit would be taken later on.
2. An *operational* discretization using fundamental measurement limitations imposed by the uncertainty relations.
3. A discretization using a *physical cutoff* imposed (e.g. by gravity).

The first steps towards a field theory over a discrete space—along the first motivation (in the context of field theory)—were taken by Ambarzumian and Iwanenko in 1930. This paper also includes a discussion of whether time would need to be quantized, along with space, as a corollary. The argument is simple: a minimum length implies a maximum frequency which implies a minimum time interval $\Delta t = \frac{1}{c}\Delta x$.[10] They are concerned solely with the infinite self-energies that arise from the point-like nature of electrons. As

[9] As Rueger makes clear (Rueger 1992, 317), prior to the 1930s there was a sense that the infinities were simply a hangover from the classical theory that if cured first (classically) would not reassert themselves at the quantum level. This was not the case, and it became clear that there existed specifically quantum divergences.

[10] In another paper from the following year, Iwanenko reiterates that the value $\lambda \sim \frac{h}{mc}$ also determines a "chronon": "Dieser Wert hat schon als kleinste definierbare Entfernung zu gelten und nicht der Elektronenradins. Mit der kleinsten Entfernung hängt die kleinste Zeitspanne zusammen" (Ivanenko 1931, 623). As Kragh and Carazza note, there were earlier speculations, with similar results, about time atoms from Pokrowski and Fürth (Kragh and Carazza 1994, 457–458). Indeed, they show that the 20s and 30s were positively teeming with discrete space, time, and spacetime proposals. However, many of them are detached from the central problems of field theories that concern us here.

they note, there seem to be two broad ways out of the predicament: give the electrons a finite size, or else restrict the spatial resolution to which one can probe (placing a limit on the validity of the theory—motivation two above). Since the former was thought to be not possible in quantum mechanics, they opt for the latter strategy. They resolve this "problem of space" by introducing a cubic lattice with grid points separated by some constant factor, a, to be determined (such that ordinary quantum theory is recovered as $a \to 0$). Differential equations are then replaced by discrete, difference equations.

This was followed by Heisenberg,[11] Ruark, March and several others, including, in England, Henry Flint. Flint was an interesting case, since he had his eye on the problem of unification of relativity and quantum mechanics in his work on fundamental length (via "ultimate measurements"—again, corresponding to the second motivation).[12] The Ambarzumian and Iwanenko paper was also directly cited by Schild, in his paper on discrete spacetime (included in this volume).

There were some other interesting attempts for "quantizing space" in the 30s. The most interesting is perhaps John Von Neumann's (unpublished) proposal from 1937.[13] Von Neumann distinguishes two kinds of singularity: the point-particle singularity and the infinite degree of freedom singularity (resulting from the infinite number of parameters needed to describe a field). In a letter to Rudolf Ortvay from 1938 he describes his model for discrete spacetime as follows:

> *(1)* The x, y, z coordinates and the t are *non-commuting* operators.
>
> *(2)* The order of magnitude of commutators is $\frac{h}{mc}$. (That is to say, this is the uncertainty associated with a simultaneous measurement of coordinates.)
>
> *(3)* The whole structure has the Lorentz-symmetry.
>
> *(4)* Each of the x, y, z coordinates has a discrete spectrum: $\pm 1/2$, $\pm 3/2$, ...
>
> *(5)* The spectrum of the time t is continuous, from $-\infty$ to $+\infty$.
>
> *(6)* When 4. and 5. are combined with 3. this comes out:
> Given four real numbers α, β, γ, δ, the spectrum of the operator $\alpha x + \beta y + \gamma z + \delta t$ is as follows:
>
> *(a)* If $\alpha^2 + \beta^2 + \gamma^2 - \delta^2 > 0$ then it is discrete: $\pm\epsilon/2$, $\pm 3\epsilon/2$, ..., where $\epsilon = \sqrt{\alpha^2 + \beta^2 + \gamma^2 - \delta^2}$.
>
> *(b)* If $\alpha^2+\beta^2+\gamma^2-\delta^2 < 0$ (indeed even when $= 0$) then it is continuous from $-\infty$ to $+\infty$.
>
> So this a "discrete" crystalline space with "continuous" time, which has not only spherical symmetry (even though it is a "crystal"!), but is even invariant

[11] In fact, in the acknowledgements to their paper, Ambarzumian and Iwanenko refer to analogous work of Heisenberg that they had only just became aware of at the time of publication.

[12] In his notebook (from 1950) there is a section on "The Theory of Relativity and the Quantum Theory" in which he nails down his project: "The underlying theme of this work is the union which exists between the theory of relativity and the quantum theory and the purpose is to portray it by means of geometry and a theory of measurement" (Henry Flint Papers, University of London, Document B53: p. 1).

[13] The manuscript is entitled "Quantum Mechanics of Infinite Systems" see Rédei (2005, 21–22).

> with respect to changes of the reference system given by Lorentz transformations, and so shows the proper Lorentz-FitzGerald contraction phenomena. (This is made possible, of course, by the non-commuting nature of the coordinates.) (von Neumann, letter to Rudolf Ortvay: March 17, 1938 [Rédei 2005, 22])

In a letter to Dirac (dated January 27, 1934) he writes:

> It should be perhaps desirable, to have operators X, Y, Z which gave discrete (point) spectra, in order to avoid the difficulties connected with the point electron (in electrodynamics). (Rédei 2005, 21)

Dirac replied the following month (February 28, 1934) pointing out that the model was not invariant under displacement of the origin of the reference frame defined by the X, Y, Z.

Quantization here is viewed, then, as a cutoff to prevent the ability to resolve to point-like distances. The problem with such accounts is that they are physically *ad hoc* (motivation one from above). Von Neumann did not pursue the idea further for this reason: "because [he] considered it very artificial and arbitrary" (Rédei 2005, 22).

Heisenberg was inspired primarily by the second motivation, though it mixed with the first, in order to tame the infinite self-energy of electrons.[14] His first thoughts about discretisation can be found in a long letter to Bohr from March 1930 (translated into English in Carazza and Kragh (1995), along with a reconstruction of the logic of the argument it contains)—one wonders whether he was aware of Ambarzumian's and Iwanenko's work, which is remarkably similar (as mentioned, Ambarzumian and Iwanenko note, at the proofs stage of their paper, that they were aware of Heisenberg's attempt, though it is hard to discern whether their work was initially written without knowledge of this). The idea is also to divide space up into a cubic lattice, where the cells have volume $r_0^3 = (h/Mc)^3$. The length $\sqrt[3]{r_0}$ (the electron radius) was then the "elementary length". He called the world described by this theory "gitterwelt" ("lattice world"). The self-energy of an electron would be rendered finite in the gitterwelt—a point Heisenberg returned to in his paper "Die Selbstenergie des Elektrons" (submitted in August of that year). As Heisenberg also notes, in the given scheme differential equations would have to be replaced with difference equations.[15] A central problem, as Heisenberg saw it (and as would deter others from the discrete space idea) was that relativistic invariance was spoiled by any scheme that introduced a fundamental length—this assumption was progressively taken apart in papers from the late 1930s onwards. (Heisenberg also pointed to difficulties in making the space isotropic; as well as with energy, momentum, and charge conservation: for these reasons he asked Bohr whether he thought the idea "completely mad"!) But beyond this breakdown of Lorentz invariance, the other target of Heisenberg's 1930 paper was to show that there are wider problems with field theory that go beyond the problem of infinite self-energy—this became part of a general programme of getting clearer on the distinct kinds

[14] Interestingly, Heisenberg had already briefly considered the idea of letting spatial coordinates be non-commuting in 1930 in order to generate a minimum length from uncertainty relations. He put this idea to Rudolf Peierls asking for any suggestions, including any input from Pauli. Julius Wess (2001, 1) claims that Heisenberg relayed it to Peierls (his student), who relayed the idea to Pauli who relayed it to Oppenheimer (his student), who relayed it to Hartland Snyder (his student: see below)! This occurs over a period of 15–16 years.

[15] Carazza and Kragh (1995) argue that Heisenberg did not really endorse a discrete space at this stage, but rather used discreteness only at the level of derivatives with respect to spatial coordinates (which are indeed replaced by discrete, finite differences).

of divergences in physical theories (on which, see the introduction to the previous part of this volume).

In his 1938 paper (Chapter 26: "Über die in der Theorie der Elementarteilchen auftretende universelle Länge"), Heisenberg explicitly ignores gravitational interactions "which hardly play a role in nuclear physics," focusing on $\hbar$ and c alone. This is part of a general to and fro with respect to the role of gravitation in elementary particle physics. However, when discussing the "universal length" he does briefly return to the issue, though again to dismiss gravity's role in the fundamental length. As is standard, he considers the electromagnetic analogy, comparing the gravitational interaction of photons with the electrical interaction of electrons. But he notes a crucial dis-analogy: introducing the gravitational constant (Heisenberg uses γ rather than G) together with $\hbar$ and c can be combined to generate the (Planck) length: $l = \sqrt{\hbar\gamma/c^3}$ (which Heisenberg computes to be 4×10^{-33}cm). However, given the vast distances separating these domains, Heisenberg points out that the problems ass'ociated with his r_0 (the electron radius) ought to be resolved first, as the most urgent task. In other words, there is a practical argument here for the *neglect* of issues having to do with quantum gravity.

The reason for this urgency were the difficulties faced by Fermi's theory of β decay, based on Pauli's neutrino hypothesis, which was found to suffer from divergences of an extreme (i.e. unrenormalizable in modern parlance) kind—involving the divergence of (Born approximation) cross sections as the energy of the incident particles went to infinity—so that the perturbation technique for treating interactions didn't give sensible answers.[16] Of course, we know that this problem was pointing to a limit with the then current quantum field theory. But Heisenberg, viewing Fermi's theory as a fundamental (and unified, in terms of weak and strong forces, with a single coupling constant) theory, took it to point to another source in which one could only resolve distances to certain distances, again close to his r_0.[17] In this case Heisenberg drew attention to the particle multiplicity ("explosionen") in cosmic ray showers in which many particles are created: the particle production would limit the resolution (so that r_0 represents a fundamental limit in this sense: physics becomes "turbulent" at shorter lengths as the coupling blows up). Of course, there was a limit, but the limit was theoretical rather than practical: there was a layer of particle physics below that captured by Fermi's theory. The short-distance, strong interaction physics that followed this was a major impetus to quantum gravity physics since gravitational and strong interactions had similar non-linearities (due to the self-interacting nature of the forces)—though, of course, gravity is universal (couples to all sources of energy equally). Hence, a new analogy between these forces, and to a lesser extent with electrodynamics, took hold.[18]

[16]The is the famous four-fermion coupling G_F which was not properly understood until the electroweak theory was developed, and the machinery of gauge theory was applied, along with Yukawa's idea of mediation by a new kind of boson (the "U-quantum" or mesotron) which replaced the four-fermion term. The evidence for mesons came in 1937, when they were isolated in cosmic rays. It should perhaps also be said that the realisation that all would not be plain sailing with respect to the other forces of nature shifted the focus onto the peculiarities of gravitation. See Chapters 3 and 4 of Brown and Rechenberg (1996) for a historical study of the Fermi-field theory.

[17]Now the length involves the mesotron mass μ, $\hbar/\mu c$, derived from Yukawa's theory. Note that Heisenberg's persistent belief in a fundamental, universal length can be seen as more reasonable given that there is a remarkable coincidence between the electron radius and this meson mass (and so the range of the nuclear forces).

[18]As Brown and Rechenberg make clear, the existence of cosmic ray phenomena was pivotal precisely in that it served to delineate the borders of the known physical theories, pointing out exactly when they would

We have seen that Bronstein had already written on related issues in 1936, pointing out that there exist quantum measurement restrictions beyond the commutation relations in the case of gravitational measurements, since there cannot be bodies of arbitrarily large mass density (cf. Gorelik and Frenkel 1994, 106). Bronstein thought this called for a revision of spacetime concepts (as did Heisenberg, though for different reasons). Unfortunately, his untimely death means that we don't know how or if he intended to pursue this revision.

The third motivation was discussed, mostly informally, in our pre-1950 period, by Pauli and Landau amongst others, but the idea was not fully developed in published form beyond mere suggestions.

Hartland Snyder's is probably the best known early work on discrete spacetime—the paper is truly a citation classic, with 1714 citations at the time of writing.[19] In this case the resulting spacetime is explicitly presented as *quantized*, with the spacetime coordinates themselves represented by Hermitian operators with discrete spectra. We have already seen this basic idea, of course, with Heisenberg and von Neumann. The innovation is to properly formalise the idea and establish that the discrete space idea need not violate Lorentz invariance.

Snyder (1947) returned to the subject the following year, with a paper applying the quantized spacetime concept to the electromagnetic field. There is then a trail leading from Snyder to Schild, in which the mathematical properties of quantized spacetime are uncovered.[20]

C. N. Yang (Chapter 28) tackled a serious issue with Snyder's model, namely that it violates translation invariance whenever the coordinates are not a continuum. However, a continuum clashes with the fundamental (i.e. non-epistemic) minimum length of the model.[21] Yang resolves the translation issue, but a problem of scale persists, namely in the form of a curvature of the universe at odds with what we observe (curvatures are of the order of the Planck scale rather than the Hubble radius).

---

break down, see Brown and Rechenberg (1996, 72). Heisenberg was, of course, wrong in thinking that Fermi's theory was fundamental: there was new physics that Heisenberg was not then privy to.

[19]Of course, this is the same Snyder who had worked with Robert Oppenheimer, in 1939, on the fate of very massive collapsing stars (approximated by an homogeneous, zero pressure ball of dust), showing that a one-way membrane (an "event horizon" in modern parlance) would emerge from the process and that a final singularity would also result—Landau had earlier noted the existence of a critical mass in 1932, and Chandrasekhar had shown in 1931 that the electron degeneracy pressure could not withstand further collapse for stars greater than 1.3 solar masses. It is rather odd that Snyder never made any link between these two streams of his work—continued collapse to a singularity and discrete space—since the former involves the reduction of a system's dimensions to values small enough (perhaps indefinitely small) to be relevant for the latter. (Oppenheimer and Snyder even write, "Physically such a singularity would mean that the expression used for the energy-momentum tensor does not take account of some essential physical fact which would really smooth the singularity out", Oppenheimer and Snyder (1939, 456). Later, John Wheeler would bring the two together via the Planck length: two areas where the "dynamics of geometry" fails to lend itself to classical analysis (Wheeler 1968, 253–254). This work would lead, ultimately, to Wheeler's notion of "spacetime foam."

[20]Bergmann and Brunings briefly refer to Snyder, if only to distance their quantised metric variables from his: their coordinates, as they say, "commute with each other, but not with the energy-momentum densities". They continue: "The dynamical character of any particle coordinates follows automatically, but probably does not exhaust the physical significance of the coordinate commutation relations" (see Chapter 33). This highlights the continuity, at least, between Snyder's (and the other related) work on quantized/discrete spaces and quantum gravity research.

[21]We saw above, in footnote 8, how Rovelli and Speziale manage to sidestep the problem by introducing probabilities for measurement outcomes.

It is rather interesting that Alfred Schild published his work on discrete spacetime around the same time that he transitioned into research on the canonical quantization of general relativity, following Dirac's influential 1949 lectures at the International Mathematical Congress in Canada, which Schild attended with his Masters student Felix Pirani.[22] Yet there is no mention of gravity in his paper on discrete spacetime, despite the fact that together with Pirani, Schild explicitly quantised the spacetime metric. This clearly reveals (perhaps rather surprisingly) that the project of "quantization of gravity" had not yet been linked to what we now call "quantum spacetime." The focus is instead on the construction of a discrete model of spacetime that is as close as possible to Lorentz invariant, and the context is the problematic divergences of standard quantum field theory. Schild's basic object is a hypercubic lattice, with time coordinate included amongst the spatial coordinates. He deals with ($c = 1$) Lorentz transformations that map a 3-lattice onto itself (where the 3-lattice takes on integer coordinates).

While Snyder's approach was indeed Lorentz invariant, it made use of the rather awkward idea that spacetime coordinates were non-commuting operators (so that spacetime functions become Hilbert space operators) and was not translation-invariant. Schild uses coordinates that are integer multiples of a fundamental length (rather than having eigenvalues that are integer multiples as with Synder), and so more along the lines of the proposals of Ambarzumian and Iwanenko *et al.*. Schild's goal is likewise to show that a common objection against discrete approaches to eliminating the divergences—that they violate Lorentz invariance due to the frame dependence of the "minimum" cell size—is only partially correct since one can construct models that are invariant under a large subgroup (the discrete subgroup) of the Lorentz group. These, he suggests are in fact physically viable (unlike Snyder's and Yang's), and cast in a model closer to ordinary spacetime, thus undermining a host of common objections and making discrete models in principle a genuine possibility for fundamental physical theory—though, as he admits, his own model suffers from physical inconsistencies to do with a radically oversized minimal velocity.[23]

Nathan Rosen (Chapter 29) introduces statistical considerations into the treatment of a discrete space: his elementary volumes are related to position measurement uncertainties (that is, to *practical* limitations: no infinitesimal measurement rods, therefore no physical point-like measurements). More specifically, the measurement of spatial coordinates of elementary particles (electrons) introduces inaccuracy into the measurement results such that repeated measurements will generate values sitting around the mean of a Gaussian distribution. His aim is, as with other proposals we've considered, to eliminate singularities (relating to the second motivation again). The resulting picture is not so very different from the Synderian one of a non-commutative space. However, the discreteness here is epistemological, coming from the difficulties involved in pinning down a spacetime point.

There is a very (later) Eddingtonian quality to this, especially the splitting of the abstract space from the observable space, which corresponds to Eddington's geometrical and

[22]Indeed, Schild's paper appears in the very same journal as Dirac's paper, in the issue directly preceding that containing the paper that would inspire Schild's work on the quantisation of the gravitational field.

[23]This shortcoming was partially eliminated by E. L. Hill in 1955 by restricting the values of spacetime variables to rational numbers—partially, because the resulting space does not quite live up to the "discrete" moniker. As Hill notes in a footnote in this paper, his Master's student, C. N. Kelber was working on this same problem of Lorentz invariance violation at the same time as Schild. There is some correspondence between Schild and Kelber, where the latter explains that he has a model that involves non-homogeneous Lorentz transformations so that the origin is not fixed for all observers (Kelber, letter to Schild, June 21st 1948).

physical frames.[24] Volumes in the observable space correspond to points in the abstract space. Lorentz invariance is preserved in this scheme only in the abstract space; yet Rosen suggests that a kind of translation manual could be established between transformations in this space and real physical transformations in the observable space.

There is also an interesting parallel here to some of the issues regarding the "reality of spacetime points" (e.g. in the context of the hole argument in general relativity). Rosen argues that the value of a physical quantity at a point is not directly observable, so that physical laws should not be based on such quantities. What is not clear is whether, according to Rosen, the world (ontology) tracks epistemology so that our laws must be written this way because the world is that way so that only the mean values of quantities over volumes have any physical meaning at all.

Rosen reviewed a closely related paper by Averbah and Medvedev in 1949. He also later returned to a similar idea, writing with Asher Peres, in 1960, though this time explicitly linking to measurement of the gravitational field. By this stage they viewed the existence of quantum uncertainties in these measurements (in the mean values of the Christoffel symbols) as pointing to the necessity of quantizing the gravitational field. Though we don't see any explicit discussion of the "discrete space-gravitation" connection, the works presented here nonetheless contain crucial evolutionary steps. The recognition that playing around with the structure of space(time) might offer up cures for some of the difficulties of quantum field theory was an early one; linking this up with the way in which general relativity includes the geometrical structure of spacetime as one of the dynamical variables took somewhat longer.

References

Brown, Laurie M. (1993). *Renormalization: From Lorentz to Landau (and Beyond)*. Berlin: Springer.

Brown, Laurie M. and Helmut Rechenberg (1996). *The Origin of the Concept of Nuclear Forces*. Bristol and Philadelphia: Institute of Physics Publishing.

Carazza, Bruno and Helge Kragh (1995). Heisenberg's Lattice World: The 1930 Theory Sketch. *American Journal of Physics* 63(7):595–605.

Cushing, James T. (1990). *Theory Construction and Selection in Modern Physics: The S Matrix*. Cambridge: Cambridge University Press.

Dantzig, David van (1938). Some Possibilities for the Future Development of the Notions of Space and Time. *Erkenntnis* 7:142–146.

— (1956). On the Relation Between Geometry and Physics and the Concept of Space-Time. In: *Jubilee of Relativity Theory: Proceedings*. Ed. by André Mercier and Michel Kervaire. Basel: Birkhäuser, 48–53.

Deser, Stanley (1957). General Relativity and the Divergence Problem in Quantum Field Theory. *Reviews of Modern Physics* 29(3):417–423.

DeWitt, Bryce (1962). The Quantization of Geometry. In: *Gravitation: An Introduction to Current Research*. Ed. by Louis Witten. New York: Wiley. Chap. 8, 266–381.

— (1964). Gravity: A Universal Regulator? *Physical Review Letters* 13(3):114–118.

Gorelik, Gennady E. and Victor Ya. Frenkel (1994). *Matvei Petrovich Bronstein and Soviet Theoretical Physics in the Thirties*. Basel: Birkhäuser.

[24] Rosen adds a statement to this effect, though pointing out that he was unaware of Eddington's work at the time of writing. Interestingly, this information was relayed to Rosen by M. F. M. Osborne, more famous now perhaps as an early econophysicist, but who also did early work on measurement restrictions, and minimal length, in quantum general relativity.

Hagar, Amit (2014). *Discrete or Continuous? The Quest for Fundamental Length in Modern Physics*. Cambridge: Cambridge University Press.
Ivanenko, Dimitri (1931). Die Beobachtbarkeit in der Diracschen Theorie. *Zeitschrift für Physik* 72(9–10):621–624.
Klein, Oskar (1956). Generalisations of Einstein's Theory of Gravitation Considered from the Point of View of Quantum Field Theory. In: *Jubilee of Relativity Theory: Proceedings*. Ed. by André Mercier and Michel Kervaire. Basel: Birkhäuser, 58–71.
Kragh, Helge and Bruno Carazza (1994). From Time Atoms to Space-Time Quantization: The Idea of Discrete Time, ca 1925–1936. *Studies in History and Philosophy of Science* 25(3):437–462.
Landau, Lev D. (1955). On the Quantum Theory of Fields. In: *Niels Bohr and the Development of Physics*. Ed. by Wolfgang Pauli. London: Pergamon Press.
Major, Seth, David Rideout, and Sumati Surya (2007). On Recovering Continuum Topology from a Causal Set. *Journal of Mathematical Physics* 48(032501).
Mercier, André and Michel Kervaire, eds. (1956). *Jubilee of Relativity Theory: Proceedings*. Basel: Birkhäuser.
Meyenn, Karl von, ed. (2001). *Wolfgang Pauli: Wissenschaftlicher Briefwechsel mit Bohr, Einstein, Heisenberg u.a.* Vol. 4, Part 3: 1955–1956. Berlin: Springer.
Oppenheimer, J. Robert and Hartland Snyder (1939). On Continued Gravitational Contraction. *Physical Review* 56:455–459.
Rédei, Miklós, ed. (2005). *John von Neumann: Selected Letters*. Providence: American Mathematical Society.
Rickles, Dean (2014). *A Brief History of String Theory: From Dual Models to M-Theory*. Berlin / Heidelberg: Springer.
Rovelli, Carlo and Simone Speziale (2003). Reconcile Planck-Scale Discreteness and the Lorentz-Fitzgerald Contraction. *Physical Review D* 67(064019).
Rueger, Alexander (1992). Attitudes towards Infinities: Responses to Anomalies in Quantum Electrodynamics, 1927–1947. *Historical Studies in the Physical and Biological Sciences* 22:309–337.
Snyder, Hartland (1947). The Electromagnetic Field in Quantized Space-Time. *Physical Review D* 72:68.
Wess, Julius (2001). Nonabelian Gauge Theories on Noncommutative Spaces. *Communications in Mathematical Physics* 219(1):247–257.
Wheeler, John A. (1968). Superspace and the Nature of Quantum Geometrodynamics. In: *Battelle Rencontres: 1967 Lectures in Mathematics and Physics*. Ed. by Cécile DeWitt-Morette and John A. Wheeler. New York / Amsterdam: W. A. Benjamin, 242–307.
Wigner, Eugene P. (1956). Relativistic Invariance of the Quantum Mechanical Equation. In: *Jubilee of Relativity Theory: Proceedings*. Ed. by André Mercier and Michel Kervaire. Basel: Birkhäuser, 210–226.

Chapter 25
Victor Ambarzumian and Dmitri Iwanenko (1930): Zur Frage nach Vermeidung der unendlichen Selbstrückwirkung des Elektrons

Victor Ambarzumian and Dmitri Iwanenko (1930). Zur Frage nach Vermeidung der unendlichen Selbstrückwirkung des Elektrons. *Zeitschrift für Physik*, 64: 563–567.

Zur Frage nach Vermeidung der unendlichen Selbstrückwirkung des Elektrons.

Von **V. Ambarzumian** und **D. Iwanenko** in Charkow.

(Eingegangen am 21. Juli 1930.)

Es wird versucht, die in der Quantenelektrodynamik eintretende Schwierigkeit der unendlichen Rückwirkung des Elektrons auf sich selbst durch Einführung von Differenzengleichung anstatt Differentialgleichungen zu vermeiden. Diese Auffassung gestattet die von Klein* an einem Beispiel betonte Schwierigkeit der relativistischen Wellengleichung im wesentlichen zu beseitigen.

Die Quantenelektrodynamik von Heisenberg und Pauli** führt zu unendlicher Rückwirkung des Elektrons auf sich selbst. Der Grund dafür besteht in der in der Quantenmechanik bisher so gut bewährten Annahme des punktförmigen Elektrons. In der klassischen Theorie konnte man diese Schwierigkeit durch die Einführung eines endlichen Elektronenradius r_0 überwinden (obgleich auch nicht auf ganz einwandfreiem Wege). Eine solche Annahme ist in der Quantenmechanik nicht möglich.

Es hat nämlich überhaupt keinen Sinn, über die Struktur des Elektrons zu sprechen, da die Bestimmung derselben sich notwendig auf die Messung der Entfernungen zwischen je zwei Punkten des Elektrons zurückführen muß. Die Messung einer Strecke auf dem Elektron muß man aber z. B. mittels eines „γ-Strahlmikroskops" durchführen. Da der Radius des Elektrons sicher nicht größer als 10^{-12} cm ist, sind wir gezwungen, Strahlen mit einer Wellenlänge zumindest nicht größer als 10^{-13} cm zu benutzen. Dann wird der Elektronenradius in erster Näherung gemessen. Unter Einwirkung dieser Lichtquanten erleidet das Elektron einen Rückstoß, ändert also seine Geschwindigkeit. Die Größe der Geschwindigkeitsänderung hängt von der Richtung ab, ist aber von der Ordnung der Lichtgeschwindigkeit. Solche Unbestimmtheit in der Geschwindigkeit für den Beobachtungsmoment bedingt eine entsprechende Unbestimmtheit für die ausgerechnete Länge auf dem ruhenden Elektron, da die Lorentzkontraktion auch unbestimmt ist.

Es hat also überhaupt keinen Sinn, über die Form und Struktur des Elektrons im gewöhnlichen Sinne zu sprechen, denn der Fehler in der Bestimmung einer Länge auf dem Elektron hat die Größenordnung dieser Länge selbst. Es scheint daher nötig, den ganzen üblichen Begriff der

* O. Klein, ZS. f. Phys. **45**, 189, 1927.
** W. Heisenberg und W. Pauli, ZS. f. Phys. **61**, 1, 1929.

räumlichen Ausdehnung für solche kleinen Partikeln zu ändern. Hier möchten wir einen ganz vorläufigen Ausweg vorschlagen, der aber vielleicht dazu dienen kann, eine konsequente Theorie, die endgültig das Problem des Raumes in Quantenmechanik lösen wird, aufzusuchen.

1. Wir führen im dreidimensionalen Raume ein kubisches ganzzahliges Punktgitter mit der vorläufig noch unbestimmt bleibenden Gitterkonstante a ein.

Wir fordern, daß die Elektronen sich nur in Gitterpunkten befinden können, also nur die folgende Koordinaten haben sollen:

$$x = ka, \quad y = ma, \quad z = na \quad (k, m, n = 0, 1, 2 \cdots -1, -2 \cdots)$$

Wir müssen also alle Gleichungen der Atommechanik als Differenzengleichungen aufschreiben. Dabei sollen unsere Differenzengleichungen bei $a \to 0$ in gewöhnliche Differentialgleichungen der Quantenmechanik übergehen. Natürlich sollen dabei alle Koordinatendifferenzen endlich bleiben, also unsere Quantenzahlen k, m, n gegen ∞ streben. Die Lösungen der Differenzengleichungen müssen zu entsprechenden Lösungen von Differentialgleichungen konvergieren.

Angenähert können wir annehmen, daß jeder Operator $\partial/\partial x$ durch die entsprechende Differenzbildung $\varphi_x = \varphi(x + a) - \varphi(x)/a$ ersetzt wird. Die Form der Gleichungen bleibt dabei beibehalten*.

Unsere nächste Aufgabe ist die Berechnung der Wechselwirkung des Elektrons mit sich selbst. Dazu müssen wir die Greensche Funktion $g(p_1 p')$ der Laplaceschen Gleichung:

$$\varphi_{xx} + \varphi_{yy} + \varphi_{zz} = 0 \tag{1}$$

bei $p = p'$ aufsuchen, denn die Rückwirkung des Elektrons mit sich selbst ist nach der Heisenberg-Paulischen Quantenelektrodynamik gleich:

$$\varepsilon = 2\, e^2 g(p_1 p) \tag{2}$$

[in der Arbeit von H. P. ist nur die elektrostatische Rückwirkungsenergie ausgerechnet, vom Betrag $\frac{1}{2}\, e^2 g(p_1 p)$. Wenn man aber auch die magnetische Wechselwirkungsenergie berücksichtigt, so muß man den viermal größeren Wert nehmen, denn bei Summation nach allen Zuständen können wir nicht die Glieder der Größenordnung $(v/c)^2$ vernachlässigen].

Wir begnügen uns mit einer approximativen Ausrechnung des Wertes von $g(p_1 p)$. Nehmen wir an, daß auf einem gewissen großen Abstand die Funktion $g(p_1 p')$ den gewöhnlichen Charakter des Coulombschen Gesetzes hat, d. h. die Form $1/r_{p p'}$, so erhalten wir aus unserer Diffe-

* R. Courant, K. Friedrichs und H. Loeny, Math. Ann. **100**, 32, 1928.

renzengleichung ein System von linearen Gleichungen, welches $g(p_1 p)$ definiert. Dabei ist natürlich angenommen, daß das System in einen Kasten gelegt wird, der genügend große Abmessungen hat. Wir haben so viele Gleichungen, wie es Punkte innerhalb des Gebietes gibt, an dessen Rand wir die klassischen Werte von $g(p_1 p')$ vorgeschrieben haben.

Es wurden der Reihe nach die Gebiete mit einem, dann sieben und endlich 19 inneren Punkten genommen und die Rechnung durchgeführt. Dabei konvergierten die Werte von $g(p_1 p)$ schnell, und es ergibt sich mit hinreichender Genauigkeit:

$$g(p_1 p) = \frac{3{,}17}{a}. \tag{3}$$

Daraus folgt für die Selbstrückwirkungsenergie:

$$\varepsilon = \frac{6{,}34\, e^2}{a}. \tag{4}$$

Nehmen wir an, daß die Ruheenergie des Elektrons eine rein elektromagnetische Natur hat, d. h. $mc^2 = \varepsilon$, so erhalten wir:

$$a = \frac{6{,}34\, e^2}{m c^2}, \tag{5}$$

also der Größenordnung nach gleich dem klassischen Durchmesser des Elektrons.

Es wäre natürlich von Wichtigkeit, einen expliziten Ausdruck für das elektrostatische Wechselwirkungsgesetz zu bekommen.

2. Aus unseren Annahmen folgt unmittelbar die Existenz einer maximalen Packung des Raumes durch die Elektronen. Vorläufig sind wir dazu gezwungen, auch für die Protonen dasselbe Punktgitter anzunehmen. Dann folgt die Existenz eines maximalen möglichen Anstiegs des Potentials auf einer gegebenen Strecke der Länge a. Bringen wir nämlich zwei Schichten der Elektronen und Protonen auf eine solche Entfernung, so muß man zur Verwirklichung eines vorgeschriebenen Potentialsprungs (sagen wir z. B. $2\, mc^2$) eine bestimmte Anzahl der Partikel in jedem Flächenelement haben. Sei N die Anzahl der Partikel in der Flächeneinheit und l die Entfernung zwischen den Schichten, dann ist die Größe des Potentialsprungs gegeben durch:

$$V_2 - V_1 = 2\,\pi e^2 N l.$$

Setzen wir hier $V_2 - V_1 = 2\, mc^2$ und $l = a$, so bekommen wir für diesen Fall:

$$N = \frac{m c^2}{\pi e \varepsilon a}.$$

Aber nach (5) ist $\frac{mc^2}{\pi e^2} = \frac{6,34}{\pi a}$, und folglich:

$$N = \frac{3,17}{3,14} \cdot \frac{2}{a^2}. \tag{6}$$

Erinnern wir uns jetzt, daß gerade ein Sprung der Größe $2\,mc^2$ bei der von Klein durchgeführten relativistischen Behandlung des Durchgangs der Elektronen durch eine Potentialwand eine kritische Bedeutung hatte, daß nämlich im Falle eines größeren Sprunges die Wand vollkommen durchsichtig war.

Nun sehen wir, daß die dazu erforderliche Flächendichte der Partikel ungefähr gleich der maximal möglichen ist, welche nach unseren Vorstellungen überhaupt existieren kann (denn in jedem Gitterpunkt können sich nicht mehr als zwei Partikel mit entgegengesetzten Spinrichtungen befinden). Also gestattet die Einführung des diskontinuierlichen Raumes für diesen Fall das Kleinsche Paradoxon zu beseitigen.

Wenn der Abstand zwischen zwei Schichten größer ist als a und somit der maximale Potentialsprung größer als $2\,mc^2$, erscheinen die Verhältnisse recht kompliziert. Denn in der für den Durchgang der Barriere nötigen Zeit muß sich unsere Potentialwand erheblich zerstören infolge des Spiels der elektromagnetischen Kräfte zwischen Partikeln, welche die Potentialwand erzeugen, und infolge der Zerfließung der Wellenpakete der die Wand konstituierenden Partikel. Also muß sich die Höhe der Potentialwand vermindern.

Schluß. Die hier vorgeschlagene Quantelung ist nicht invariant gegen beliebige Drehungen und Verschiebungen der Koordinaten, aber es scheint uns die Möglichkeit nicht ausgeschlossen, auf eine spezielle Weise die Invarianz der Gleichungen zu sichern. Vielleicht muß man das vierdimensionale Volumen als einfachste invariante Größe einführen und fordern, daß es immer eine bestimmte Anzahl der Gitterpunkte enthalte*.

Ein anderer geistreicher Vorschlag, die Transformationen des Gitterraumes zu behandeln, rührt von Herrn Dr. H. D. Ursell her, der eine Analogie mit dem Spinelektron ins Auge gefaßt hat. Man kann nämlich festsetzen, daß ein bestimmter Punkt des Gitters nach der Transformation eine bestimmte Wahrscheinlichkeit hat, in irgendwelchem Punkte des zweiten

---

* Dieser Gedanke war von Jordan ausgesprochen, der seinerseits zu einigen analogen Ideen über die Quantelung des Raumes gekommen war. Herrn Professor P. Jordan sind wir für die kritische Bemerkung zu diesen Überlegungen zu bestem Dank verpflichtet.

Gitters getroffen zu werden (so daß eine Wahrscheinlichkeit besteht, eine andere ganzzahlige Koordinate zu bekommen). Zwar lassen sich auf diese Weise Translationen streng behandeln, die Drehungen aber bieten größere Schwierigkeiten.

Es entsteht natürlich die Frage, ob die Zeit auch gequantelt sein solle. Die Antwort scheint notwendig positiv zu lauten. Ganz abgesehen von der notwendigen vierdimensionalen Symmetrie legt die Existenz einer minimalen Entfernung auch die Annahme einer *minimalen Wellenlänge* nahe, sowohl für das Licht als auch für die Materie; und somit kommen wir zur Existenz einer *maximalen Frequenz*. Und die maximale Frequenz des Lichtes bedeutet wahrscheinlich nichts anderes als das minimale Zeitintervall, sagen wir $\Delta t = \frac{1}{c} \Delta x$. Es möge nun versuchsweise diese maximale Lichtfrequenz der Größenordnung nach der Frequenz der Lichtquanten, die bei der Vernichtung der Protonen und Elektronen entstehen, gleichgesetzt werden. Es sei also

$$\frac{\Delta x}{\Delta t} = c \quad \text{und} \quad \nu_{\max} = \frac{M + m}{2h} c^2 = \frac{1}{\Delta t}.$$

Daraus folgt: $\frac{M}{m} \cdot \frac{e^2}{hc} \approx 1$.

Die empirische Tatsache der übereinstimmenden Größenordnung von Feinstrukturkonstante und Massenverhältnis der Protonen und Elektronen kann somit theoretisch verständlich gemacht werden.

Wir möchten danach hoffen, daß die hier entworfenen Überlegungen vielleicht zur Förderung der Kernprobleme beitragen könnten.

Anmerkung bei der Korrektur. Heisenberg hat eine analoge Quantelung versucht. Es ist ihm gelungen, die Differenzenwellengleichung für das freie Elektron zu integrieren. Dabei ergab sich die sehr merkwürdige Eigenschaft des maximalen Eigenwertes. Herrn Prof. W. Heisenberg müssen wir herzlich danken für seine liebenswürdige Mitteilung.

Charkow, Physikalisch-Technisches Institut.

Chapter 26
Werner Heisenberg (1938): Über die in der Theorie der Elementarteilchen auftretende universelle Länge

Werner Heisenberg (1938). Über die in der Theorie der Elementarteilchen auftretende universelle Länge. *Annalen der Physik*, 424: 20–33.

Über die in der Theorie der Elementarteilchen auftretende universelle Länge

Von W. Heisenberg

Dieses Heft feiert den Schöpfer der Quantentheorie, die mehr als irgendeine andere Entdeckung der neueren Physik zu einer allgemeinen Veränderung und Klärung des physikalischen Weltbildes geführt hat. Diese Gelegenheit mag als Entschuldigung gelten, wenn die folgenden Überlegungen sich mehr mit allgemeinen und zum großen Teil bekannten Zusammenhängen beschäftigen, als dies vielleicht sonst im Rahmen einer Einzeluntersuchung zulässig schiene.

Vor einiger Zeit hat der Verf. darauf aufmerksam gemacht[1]), daß eine konsequente Anwendung der Fermischen Theorie des β-Zerfalls zu dem Schlusse nötigt, daß beim Zusammenstoß eines außerordentlich energiereichen Höhenstrahlungsteilchens mit einer anderen Partikel eventuell viele Teilchen explosionsartig auf einmal entstehen können. Dieser Vorgang sollte eintreten, wenn die Wellenlänge der stoßenden Teilchen im Schwerpunktsystem eine gewisse kritische Länge unterschreitet, die als eine universelle Länge von der Größenordnung des klassischen Elektronenradius

$$\frac{e^2}{m c^2} = r_0 = 2{,}81 \cdot 10^{-13}\ \mathrm{cm}$$

in die Fermische Theorie des β-Zerfalls eingeht. Die Existenz solcher Explosionen schien damals durch die Experimente über die Schauer der Höhenstrahlung und die Hoffmannschen Stöße sehr wahrscheinlich. Inzwischen haben jedoch die Theorien von Carlson und Oppenheimer[2]) sowie Bhabha und Heitler[3]) gezeigt, daß ein großer Teil dieser Phänomene einfach auf Grund der Quantenelektrodynamik durch die sogenannte Kaskadenbildung gedeutet werden kann. Ferner sind von Yukawa[4]) und Wentzel[5]) andere Theorien des β-Zerfalls vorgeschlagen worden, die anscheinend die Experimente ebenso gut darstellen wie die Fermische Theorie, und die darüber hinaus eine Verknüpfung des β-Zerfalls mit der Existenz

1) W. Heisenberg, Ztschr. f. Phys. **101**. S. 533. 1936.
2) J. F. Carlson u. J. R. Oppenheimer, Phys. Rev. **51**. S. 220. 1937.
3) H. Bhabha u. W. Heitler, Proc. Roy. Soc. A **159**. S. 432. 1937.
4) H. Yukawa, Proc. Phys. Math. Soc. Japan **17**. S. 48. 1935.
5) G. Wentzel, Ztschr. f. Phys. **104**. S. 34. 1937; **105**. S. 738. 1937.

der von Neddermeyer und Anderson[1]) neu entdeckten Teilchen in Aussicht stellen[2]). Daher steht die These, daß eine universelle Naturkonstante der Dimension einer Länge in dem Auftreten von Explosionen beim Zusammenstoß von Teilchen kleinerer Wellenlänge sichtbar werde, keineswegs auf sicherem Grund. Es sollen daher im folgenden die Argumente einzeln besprochen werden, welche die Vermutung nahelegen, es müsse in der Kernphysik und der Theorie der Ultrastrahlung auf eine universelle Länge der Ordnung r_0 Rücksicht genommen werden, und welche mit der möglichen Existenz der Explosionen in Zusammenhang stehen.

I. Die Konstanten c und $\hbar$

Die Entwicklung der neueren Physik lehrt, daß die Wirksamkeit einer grundlegenden Naturkonstanten zunächst durch die Widersprüche bemerkt wird, die sich beim konsequenten Ausbau von scheinbar experimentell völlig gesicherten umfassenden Theorien, insbesondere beim Zusammenfügen von zwei solchen Theorien, ergeben. So galt um die Jahrhundertwende die Newtonsche Mechanik sowie die Maxwellsche Theorie als gesicherter Besitz der Physik. Beim Versuch, beide Theorien zu vereinigen und die Elektrodynamik bewegter Körper auszuarbeiten, stellten sich jedoch schwere Widersprüche heraus, die erst behoben werden konnten, als man in der Lichtgeschwindigkeit c eine universelle Naturkonstante von allgemeinster Bedeutung erkannte, auf die bei der Formulierung jedes physikalischen Gesetzes Rücksicht genommen werden muß. In ähnlicher Weise mußte die von Gibbs und Boltzmann ausgearbeitete statistische Theorie der Wärme als endgültig angesehen werden. Bei der Anwendung dieser Theorie auf die Strahlungsprobleme, also beim Zusammenfügen der Maxwellschen Theorie und der Theorie der Wärme stellten sich jedoch sehr grobe Widersprüche heraus: für die Strahlung im Wärmegleichgewicht erhielt man im Rayleigh-Jeansschen Gesetz divergente Resultate. Erst die Plancksche Entdeckung zeigte, daß man beim Zusammenfügen dieser Theorien auf eine universelle Naturkonstante von der Dimension einer Wirkung Rücksicht nehmen muß.

---

1) S. Neddermeyer u. C. P. Anderson, Phys. Rev. **51**. S. 884. 1937; vgl. auch J. C. Street u. E. C. Stevenson, Bull. Amer. Phys. Soc. **12**. S. 2, 13. 1937.

2) Vgl. insbes. J. R. Oppenheimer u. R. Serber, Phys. Rev. **51**. S. 1113. 1937; E. C. G. Stückelberg, Phys. Rev. **52**. S. 42. 1937; H. Yukawa u. S. Sakata, Proc. Phys. Math. Soc. Japan **19**. S. 1084. 1937; N. Kemmer, Nature **141**. S. 116. 1938; H. Bhabha, Nature **141**. S. 117. 1938.

Als die allgemeinen Gesetzmäßigkeiten, die mit den Konstanten c und $\hbar$ verbunden sind, klar verstanden waren, erkannte man, daß es sich nicht eigentlich um eine Korrektur der früher als gesichert betrachteten Theorien gehandelt hatte. Die früheren Theorien haben als die anschaulichen Grenzfälle, in denen die Lichtgeschwindigkeit als sehr groß und das Wirkungsquantum als sehr klein betrachtet werden kann, weiter Bestand. Die Konstanten c und $\hbar$ bezeichnen vielmehr die Grenzen, in deren Nähe unsere anschaulichen Begriffe nicht mehr ohne Bedenken verwendet werden können. Man hat diesen Sachverhalt oft durch die Behauptung ausgedrückt, die früheren Theorien gingen aus Relativitätstheorie und Quantentheorie durch den Grenzübergang $c \longrightarrow \infty$, $\hbar \longrightarrow 0$ hervor. Diese Formulierung ist aber deswegen nicht ganz unbedenklich, weil sie nur dann richtig sein kann, wenn bestimmte Größen bei diesem Grenzübergang konstant gehalten werden (z. B. beim Übergang von der Quantenmechanik zur klassischen Mechanik die Massen und Ladungen der Elementarteilchen). In einer endgültigen Theorie aber würden diese Größen selbst aus den wenigen universellen Konstanten der Physik bestimmt sein, und eine Größenänderung der universellen Konstanten könnte an der Form der physikalischen Gesetze überhaupt nichts ändern. Der bei Konstanthaltung der genannten Größen vollzogene entgegengesetzte Grenzübergang $\hbar \longrightarrow \infty$ oder $c \longrightarrow 0$ führt übrigens zu sinnlosen Resultaten. Es erscheint also richtiger, bei der ersten Formulierung zu bleiben und $\hbar$ und c einfach als die Grenzen zu bezeichnen, die der Verwendung anschaulicher Begriffe gesetzt sind[1]).

Durch diese Eigenschaft unterscheiden sich übrigens die Konstanten $\hbar$ und c von anderen, weniger fundamentalen universellen Konstanten. Z. B. hängt etwa die Boltzmannsche Konstante k mit der willkürlichen Festsetzung der Temperaturskala zusammen und könnte aus einer Theorie der Aggregatzustände des Wassers berechnet werden, oder bei Messung der Temperatur im Energiemaß ganz wegfallen. Auch eine universelle Konstante wie die Masse des Protons hat keineswegs eine so prinzipielle Bedeutung wie $\hbar$ oder c; denn der Massenbegriff hat zweifellos auch bei noch kleineren Massen einen einfach angebbaren Sinn; z. B. kann die Masse eines Lichtquants noch am Lichtdruck gemessen werden. Es scheint daher zweckmäßig, Konstanten von solch grundlegenden Eigenschaften wie $\hbar$ und c etwa als „universelle Konstanten erster Art" vor den anderen auszuzeichnen.

1) Vgl. hierzu insbes. N. Bohr, Atomtheorie und Naturbeschreibung, Berlin, Springer 1931.

Diese universellen Konstanten erster Art sind, wenn man von den beiden bisher bekannten Beispielen $\hbar$ und c verallgemeinern darf, jeweils mit einer sehr allgemeinen Eigenschaft der Naturgesetze von dem Typus einer Invarianzeigenschaft verknüpft; sie stellen allgemeine Forderungen an die Form irgendeines physikalischen Gesetzes. Die spezielle Relativitätstheorie fordert die Invarianz aller physikalischen Gesetze gegenüber der Lorentztransformation, die Quantenmechanik fordert die Vertauschungsrelationen zwischen kanonisch konjugierten Variabeln, die Existenz von Wahrscheinlichkeitsamplituden und die Invarianz der Gesetze gegenüber Drehungen im Hilbertschen Raum.

Wenn man von den Gravitationserscheinungen absieht, die in der Atomphysik kaum eine Rolle zu spielen scheinen, so sind außer $\hbar$ und c in der Mikrophysik einstweilen keine weiteren universellen Konstanten erster Art bekannt.

II. Die universelle Länge

Nun ist es aber von vornherein klar, daß es in der Atomphysik noch eine weitere „universelle Konstante erster Art" von der Dimension einer Länge oder einer Masse geben muß. Denn da sich aus den Konstanten $\hbar$ und c dimensionsmäßig keine Länge oder Masse bilden läßt, so muß die Masse der Elementarteilchen und die Dimension der Atome und Atomkerne durch eine weitere universelle Konstante festgelegt sein. Obgleich nun aus einer universellen Länge r_0 stets eine universelle Masse m und umgekehrt aus der Masse eine Länge durch die Relation $m = \frac{\hbar}{c\, r_0}$ gebildet werden kann und es insofern unwesentlich scheint, ob man von einer universellen Länge oder Masse spricht, so kommt doch wohl die physikalische Bedeutung dieser Konstante klarer zum Ausdruck, wenn man sie als eine universelle Länge einführt. Denn dann bedeutet sie wieder eine Grenze für die Anwendung anschaulicher Begriffe: der Begriff einer Länge kann nur für Entfernungen, die groß sind gegen die universelle Länge, ohne Einschränkung angewendet werden. Eine ähnliche physikalische Deutung einer universellen Masse dagegen ist nicht in einfacher Weise möglich. Denn größere und kleinere Massen sind der Messung mit jeder Genauigkeit zugänglich. Wir wollen daher die neue Konstante als Länge einführen und ihr den Wert $r_0 = \frac{e^2}{m c^2} = 2{,}81 \cdot 10^{-13}$ cm geben. Diese letztere Festsetzung ist ja offenbar innerhalb gewisser Grenzen willkürlich, ebenso wie man willkürlich als universelle Konstante der Wirkung h oder $\hbar$ einführen kann. Erst die fertige Theorie lehrt, welches die zweckmäßigste Festsetzung ist, und es ist freilich nicht wahrscheinlich, daß sich gerade der Wert r_0 als der zweckmäßigste erweisen wird; doch hat er, wie die weiteren Überlegungen zeigen werden, jedenfalls die richtige Größenordnung. Die Festsetzung über r_0 soll auch *nicht* bedeuten, daß die universelle Länge mit der Frage nach der Elektronenladung in unmittelbarem Zusammenhange stünde.

Es soll nun im folgenden auseinandergesetzt werden, daß — wie wohl schon verschiedentlich ausgesprochen wurde — die Widersprüche, die bisher in der Quantenelektrodynamik, der Theorie des β-Zerfalls und der Kernkräfte allenthalben auftraten, verschwinden, wenn man auf die Einschränkungen achtet, die durch die universelle Länge r_0 vorgeschrieben werden; daß ferner die Länge r_0 in der Theorie der Elementarteilchen eine entscheidende Rolle spielen muß, und daß die Einschränkungen der Messungsmöglichkeiten, die durch r_0 bedingt sind, vielleicht durch die Existenz der Explosionen einfach verständlich gemacht werden können.

III. Die Divergenzen

Wendet man die Vorschriften der Quantentheorie auf eine relativistisch invariante Wellentheorie an, in der auch Wechselwirkungen der Wellen (d. h. nichtlineare Glieder in der Wellengleichung) vorkommen, so erhält man, wie vielfach bemerkt worden ist, divergente Resultate. Es liegt dies daran, daß die relativistische Invarianz eine „Nahewirkungstheorie" fordert, in der die Wechselwirkung dadurch bedingt ist, daß die Fortpflanzungsgeschwindigkeit einer Welle an einem Punkte durch die Amplitude einer anderen Welle an diesem Punkt bestimmt wird. Wegen der unendlich vielen Freiheitsgrade des Kontinuums, d. h. wegen der Möglichkeit von Wellen beliebig kleiner Wellenlänge werden aber die Eigenwerte einer Wellenamplitude an einem bestimmten Punkte unendlich. Dieser Widerspruch — der ja viel Ähnlichkeit mit dem Widerspruch im Rayleigh-Jeansschen Gesetz hat — bedeutet nun offenbar nicht eigentlich, daß die relativistische Wellentheorie oder die Quantentheorie falsch und zu verbessern wären, sondern weist darauf hin, daß beim Zusammenfügen der Quantentheorie und der relativistischen Wellentheorie auf eine universelle Konstante von der Dimension einer Länge Rücksicht genommen werden muß. In der Tat haben sich viele Autoren damit geholfen, daß sie die divergenten Integrale bei einer Länge von der Größenordnung r_0 (oder bei den entsprechenden Impulsen) künstlich konvergent machten oder abbrachen, womit sich vernünftige Ergebnisse erzielen ließen. Aber dieses Abbrechen kann im allgemeinen nicht in relativistisch invarianter Weise durchgeführt werden und ist natürlich nur als sehr vorläufiger Notbehelf zu betrachten. Denn in einer endgültigen Theorie müßten statt dessen die qualitativ neuen physikalischen Phänomene, die bei Längen der Ordnung r_0 auftreten (und die selbstverständlich der relativistischen Invarianzforderung genügen) richtig berücksichtigt werden, was von selbst zur Konvergenz der Integrale führen würde.

Eine Sonderstellung nimmt in dieser Betrachtung die Frage nach der Selbstenergie des Elektrons ein, die ja schon in der klassischen Theorie als Beweis für den endlichen Elektronenradius angesehen wurde. Diese bekannte klassische Begründung für den Radius r_0 kann zweifellos nicht in die Quantentheorie übernommen werden. Denn da die Ladung des Elektrons kleiner ist als die dimensionsmäßig entsprechende Größe $\sqrt{\hbar c}$, entziehen die durch die Quantentheorie bedingten unanschaulichen Züge in der Beschreibung des Elektronenfeldes der genannten klassischen Überlegung den Boden[1]). Man hätte also vielleicht hoffen können, daß eine Quantentheorie des Elektrons und seines umgebenden Feldes existiert, in der keinerlei Selbstenergie auftritt, das Elektron also keine Ruhmasse besitzt. Eine solche Theorie, die gewissermaßen durch den Grenzübergang $r_0 \longrightarrow \infty$ aus der richtigen Theorie hervorgeht, scheint aber ebenso sinnlos wie der früher besprochene Grenzübergang $\hbar \longrightarrow \infty$ in der Quantentheorie. Sobald aber eine endliche Ruhmasse des Elektrons sich aus einer Theorie ergeben soll, so muß diese Theorie außer $\hbar$ und c noch eine universelle Länge r_0, also Elemente enthalten, die mit Elektrodynamik und Quantentheorie nichts zu tun haben. Es scheint aus diesem Grunde unwahrscheinlich, daß eine Theorie der Sommerfeldschen Feinstrukturkonstante $e^2/\hbar c$ gefunden werden kann, bevor die durch r_0 bedingten neuen Züge der Naturbeschreibung, die ja zunächst mit der Frage nach der Elektronenladung gar nicht zusammenhängen, klargestellt sind.

Bei dieser Gelegenheit kann vielleicht das Problem der Gravitationskräfte kurz gestreift werden, von denen sonst in dieser Abhandlung nicht die Rede ist. Man kann die gegenseitige Schwere zweier Lichtquanten vergleichen mit der oben besprochenen elektrischen Wechselwirkung zweier Elektronen und nach der Gravitationsselbstenergie der Lichtquanten fragen[2]). Die Tatsache, daß die Lichtquanten keine Ruhmasse besitzen, legt hier zunächst den Gedanken nahe, daß in diesem Problem vielleicht r_0 keine Rolle spielen könnte. Es stellt sich jedoch heraus, daß — im Gegensatz

1) Vgl. N. Bohr u. L. Rosenfeld, Dansk Vid. Selsk. math. phys. Medd. **12**. S. 8. 1933.

2) Vgl. L. Rosenfeld, Ztschr. f. Phys. **65**. S. 589. 1930.

zum elektrischen Analogon — die Gravitationskonstante γ zusammen mit $\hbar$ und c selbst eine Länge auszeichnet: $l = \sqrt{\frac{\hbar \gamma}{c^3}} = 4 \cdot 10^{-33}$ cm. Der Umstand, daß diese Länge wesentlich kleiner ist als r_0, gibt uns das Recht, von den durch die Gravitation bedingten unanschaulichen Zügen der Naturbeschreibung zunächst abzusehen, da sie — wenigstens in der Atomphysik — völlig untergehen in den viel gröberen unanschaulichen Zügen, die von der universellen Konstanten r_0 herrühren. Es dürfte aus diesen Gründen wohl kaum möglich sein, die elektrischen und die Gravitationserscheinungen in die übrige Physik einzuordnen, bevor die mit der Länge r_0 zusammenhängenden Probleme gelöst sind.

Die Diskussion der mit r_0 verbundenen Fragen scheint also die vordringlichste Aufgabe. Man wird zu ihrer Behandlung in erster Linie die Erscheinungen der Kernphysik und der Höhenstrahlung zu studieren haben, bei denen von den elektrischen und den Gravitationswechselwirkungen in erster Näherung abgesehen werden kann.

In diesem Gebiet der Physik ist es vor allem die Theorie des β-Zerfalls, in der sich die genannte Divergenzschwierigkeit bei der Quantelung der Wellenfelder äußert. Legt man insbesondere die Fermische Theorie des β-Zerfalls zugrunde, so führt die konsequente Anwendung der Quantentheorie zu Divergenzen von so hohem Grad, daß die Resultate nicht nur quantitativ sondern auch qualitativ von der Art abhängen, wie die divergenten Integrale künstlich in konvergente umgewandelt werden. So haben z. B. die Berechnungen von v. Weizsäcker[1]), Fierz[2]) und anderen gezeigt, daß die Kräfte zwischen den Elementarteilchen, die sich aus der Theorie des β-Zerfalls ergeben, ganz von der Art des „Abschneidens“ bei kleinen Wellenlängen abhängen. Dieser Umstand hat verschiedene Forscher veranlaßt, die qualitativen Konsequenzen der Fermischen Theorie, die in der Möglichkeit der Explosionen sich äußern, anzuzweifeln, da ja durch ein geeignetes Abschneiden bei hinreichend langen Wellen erreicht werden kann, daß die genannten Konsequenzen nicht eintreten[3]). Eine solche Schlußweise scheint mir jedoch auf einem Mißverständnis zu beruhen. Denn die Berechtigung zum Abschneiden kann doch umgekehrt nur aus den qualitativ neuen Erscheinungen entnommen werden, die bei den kritischen Wellen-

---

1) C. F. v. Weizsäcker, Ztschr. f. Phys. **102.** S. 572. 1936.

2) M. Fierz, Ztschr. f. Phys. **104.** S. 553. 1937.

3) G. Nordheim, L. W. Nordheim, J. R. Oppenheimer u. R. Serber, Phys. Rev. **51.** S. 1037. 1937.

längen eintreten. Wenn diese qualitativ neuen Erscheinungen gestrichen werden, verliert auch die Abschneidevorschrift jeden physikalischen Sinn.

Ersetzt man die Fermische Theorie des β-Zerfalls durch eine andere, in welcher der beim β-Zerfall sich abspielende Prozeß aus zwei elementaren Übergängen zusammengesetzt wird, wie dies Yukawa u. Wentzel (a. a. O.) versucht haben, so treten Divergenzen von geringerem Grad als bei Fermi auf. Die Theorie des β-Zerfalls erhält dann Ähnlichkeit mit der gewöhnlichen Strahlungstheorie, wobei an die Stelle des Lichtquants ein geladenes Teilchen mit Bosestatistik und einer Masse der Größenordnung $\hbar / c r_0$ tritt. Dieses Teilchen kann dann vielleicht mit den von Neddermeyer und Anderson (a. a. O.) vermuteten instabilen Teilchen identifiziert werden. Die Frage, ob in dieser Theorie bei den in der Höhenstrahlung vorkommenden Energien Explosionen zu erwarten sind, hängt davon ab, ob die der Sommerfeldschen Feinstrukturkonstante entsprechende Größe[1]) $g^2/\hbar c$ bei der Berechnung der Wirkungsquerschnitte als klein gegen Eins betrachtet werden kann. Ihr Wert liegt nun je nach der Masse der Yukawaschen Teilchen etwa zwischen $^1/_{10}$ und 1. Die Konsequenzen dieser Theorie für die Frage der Mehrfachprozesse sind daher einstweilen nicht zu übersehen. Im Grenzfall $\frac{g^2}{\hbar c} \ll 1$ werden die Mehrfachprozesse unwahrscheinlich; dies hat zur Folge, daß die Theorie zwar weitgehende Ähnlichkeit mit der Strahlungstheorie erhält, dafür aber die Berechtigung zu der auch in ihr notwendigen „Abschneidevorschrift" aus anderen in der Theorie nicht enthaltenen und unbekannten Phänomenen holen muß. Für Werte $\frac{g^2}{\hbar c} \sim 1$ dagegen führt auch diese Theorie zur Möglichkeit der Explosionen. Sie verliert jedoch dabei wohl die Ähnlichkeit zur Strahlungstheorie, da dann eine Entwicklung nach $g^2/\hbar c$ sinnlos werden dürfte[2]). Auf jeden Fall handelt es sich also hier — ebenso wie in der Fermischen Theorie — wohl nur um ein korrespondenzmäßiges Analogon zu einer endgültigen Theorie, in der die Länge r_0 an wesentlicher

1) H. Yukawa u. S. Sakata, a. a. O., S. 1090.

2) Wenn man die Ergebnisse von B. Kockel, Ztschr. f. Phys. **107**. S. 153. 1937 auf die Yukawa-Wentzelsche Theorie übertragen und verallgemeinern darf, so wären schon bei einem Wert $\frac{g^2}{\hbar c} = \frac{1}{10}$ die Mehrfachprozesse von etwa 10^8 eV ab die Regel. Schon für $\frac{g^2}{\hbar c} = \frac{1}{10}$ ist daher die Konvergenz einer Entwicklung nach Potenzen von $g^2/\hbar c$ sehr fraglich.

Stelle vorkommen muß. Wie fruchtbar andererseits solch ein korrespondenzmäßiges Analogon sein kann — selbst in einem Gebiet, in dem die unanschaulichen Züge bereits eine wesentliche Rolle spielen —, zeigt in der Vergangenheit etwa das Beispiel der Uhlenbeck-Goudsmitschen Theorie des Spins aufs deutlichste.

IV. Die Theorie der Elementarteilchen

Daß die bisherigen Theorien des β-Zerfalls nur den Charakter eines korrespondenzmäßigen Analogons haben können, scheint insbesondere daraus hervorzugehen, daß in ihnen die Massen der Elementarteilchen als eigene universelle Konstanten vorkommen. In der endgültigen Theorie müßten diese vielen verschiedenen Ruhmassen von Neutron, Proton, Elektron, Neutrino und den neuen labilen Teilchen sich in ähnlicher Weise aus der Konstante r_0 ergeben, wie etwa die Terme des Wasserstoffatoms aus der Rydbergkonstante. Nun wird es freilich große Bereiche der Physik geben, in denen die Massen der Elementarteilchen einfach als feste Parameter betrachtet werden können und in denen die Theorie dieser Massen auf später aufgeschoben werden kann; und zwar wird dies überall dort der Fall sein, wo es sich um Energieumsetzungen handelt, die im Schwerpunktsystem klein gegen die kritische Energie $\hbar c/r_0$ sind. In dieses Gebiet gehört nicht nur die ganze Physik der Atomhülle, sondern auch die gewöhnliche Kernphysik und die Theorie des β-Zerfalls. Erst wenn man die Theorie des β-Zerfalls zu verknüpfen sucht mit der Theorie der Kernkräfte, oder wenn man sie anwenden will auf Probleme der Höhenstrahlung, braucht man Aussagen über Prozesse mit großer Energieumsetzung. Bei solchen Aussagen wird man ein Eingehen auf die Theorie der Elementarteilchen nicht mehr vermeiden können. Man muß also schließen, daß alle Versuche, die β-Zerfallstheorie mit den Kernkräften, mit dem magnetischen Moment der Elementarteilchen und mit Prozessen der Höhenstrahlung zu verknüpfen, nur sehr vorläufigen Charakter haben können, solange in ihnen die Massen der Elementarteilchen als unabhängige Konstanten erscheinen. Denn wenn Prozesse diskutiert werden, bei denen Teilchen einer Ruhmasse der Ordnung $h c/r_0$ entstehen, wird notwendig auf die bei der Länge r_0 eintretenden qualitativ neuen Erscheinungen geachtet werden müssen. Diese Erscheinungen müssen einerseits die Elementarmassen festlegen und andererseits den Grund für die Beseitigung der in Kap. III besprochenen Divergenzen abgeben.

V. Die durch die universelle Länge r_0 bedingten neuen Erscheinungen

Welches sind nun die bei Entfernungen oder Wellenlängen der Größenordnung r_0 auftretenden neuen Erscheinungen? Solange es sich nur um die Bewegung einer einzigen Korpuskel handelt, kann wegen der relativistischen Invarianzforderungen die Konstante r_0 sich nur äußern im Auftreten einer Ruhmasse. Ob dieses einzige Teilchen eine Energie groß oder klein gegen die kritische Energie $\hbar c/r_0$ hat, ist natürlich völlig gleichgültig, da die Energie vom Bezugssystem des Beobachters abhängt.

Wenn jedoch zwei Teilchen in Wechselwirkung treten, so wird es für das weitere physikalische Geschehen wesentlich sein, ob die kinetische Energie der Teilchen im Schwerpunktsystem bei ihrem Zusammentreffen groß oder klein gegen $\hbar c/r_0$ ist. Im Falle *kleiner* Energien kann, das zeigen die Erfahrungen der Kernphysik, das Verhalten der Teilchen so aufgefaßt werden, als wirke zwischen ihnen eine Kraft, die nur auf Abständen der Ordnung r_0 (im Schwerpunktsystem) merkliche Werte annimmt. Diese Wechselwirkungsenergien der Größenordnung $\hbar c/r_0$ und der Reichweite r_0 sind sozusagen das erste charakteristische Merkmal der Konstanten r_0. Es erscheint daher auch fraglich, inwieweit es zweckmäßig ist, diese Kräfte als abgeleitet aus den β-Zerfallskräften zu betrachten. In Wirklichkeit bilden wohl die Kernkräfte und die β-Zerfallskräfte eine Einheit, und man wird kaum von primären und abgeleiteten Wirkungen sprechen können. Aus ähnlichen Gründen wird man wohl auch annehmen dürfen, daß Kräfte der ungefähren Reichweite r_0 zwischen *allen* Arten von Elementarteilchen wirksam sind; eine Ausnahme bilden höchstens die Partikeln, deren Ruhmasse sehr viel kleiner als $\hbar c/r_0$ ist (Elektronen, Neutrinos, Lichtquanten); bei denen werden vielleicht auch die „Kernkräfte" besonders schwach sein.

Sehr viel weniger als über die Wechselwirkung der Teilchen kleiner kinetischer Energie ist bekannt über die Wechselwirkung zweier Teilchen, die beim Zusammenstoß im Schwerpunktsystem eine kinetische Energie besitzen, die groß gegen $\hbar c/r_0$ ist. Offenbar müssen die Prozesse, die sich hier abspielen, eng mit den unanschaulichen Zügen zusammenhängen, die durch die Konstante r_0 in die Physik hereingebracht werden; ähnlich, wie etwa das Verhalten eines Elektrons z. B. im Normalzustand des Wasserstoffatoms die charakteristischen unanschaulichen Züge der Quantentheorie besonders deutlich zeigt.

Nun wird man freilich nicht erwarten, daß man alle Möglichkeiten für die Äußerung der von r_0 stammenden unanschaulichen Züge beim Zusammenstoß energiereicher Teilchen zu überschauen vermag.

Aber die eine Möglichkeit, die durch die Fermische Theorie nahegelegt wird, soll noch ausführlich besprochen werden.

Es kann angenommen werden, daß beim Zusammenstoß zweier Teilchen, deren Energie im Schwerpunktsystem groß gegen $\hbar c/r_0$ ist, diese Energie im allgemeinen in einem Akt in viele Elementarteilchen aufgeteilt wird. Diese Annahme der Explosionen ergibt sich als Folgerung aus der Fermischen Theorie des β-Zerfalls. Aber ganz unabhängig von dieser Theorie ist sie eine logische Möglichkeit, die allen aus Relativitätstheorie und Quantentheorie stammenden Invarianzforderungen genügt.

Die Entstehung von neuen Elementarteilchen beim Zusammenstoß zweier energiereicher Partikel wird ja schon durch die Analogie zur Elektrodynamik nahegelegt: Die Wechselwirkung energiearmer Elektronen wird durch die Coulombsche Kraft bestimmt, wie dies in den Theorien der Bremsung und Ionisation ausführlich dargestellt wird. Bei der Ablenkung sehr energiereicher Elektronen dagegen spielt nach Bethe und Heitler die Strahlung die Hauptrolle. Man kann dies so auffassen: Bei der Ablenkung des nahezu mit Lichtgeschwindigkeit bewegten Elektrons kann sein elektrisches Feld wegen der Retardierung nicht ohne weiteres folgen, ein Teil dieses Feldes verläßt als Lichtquant den Ort, wo die Ablenkung stattgefunden hat. Dabei kann, wie die Theorie von Bethe und Heitler zeigt, das Lichtquant häufig einen erheblichen Teil der Energie des abgelenkten Teilchens mitnehmen. In ähnlicher Weise kann angenommen werden, daß beim Zusammenstoß zweier Elementarteilchen, die sich mit sehr großer Energie auf Abstände der Ordnung r_0 nähern, eine Mitführung des Kernfeldes bei der Ablenkung nicht ohne weiteres möglich ist, daß also ein Teil des Feldes in Form von Elementarteilchen, die dann wieder einen großen Teil der Gesamtenergie mitnehmen können, den Ort des Zusammenstoßes verläßt.

Diese Analogie lehrt auch, daß — wenn die Explosionen überhaupt stattfinden — erwartet werden muß, daß Teilchen *aller* Art in den Explosionen entstehen können. Diese Annahme wird zwar durch die Fermische Theorie nicht nahegelegt, da in ihr z. B. die neuentdeckten labilen Teilchen nicht vorkommen; sie scheint mir jedoch eine natürliche Konsequenz aus den physikalischen Grundlagen der Hypothese der Explosionen. Insbesondere werden also bei einer Explosion häufig Neddermeyer-Andersonsche Teilchen und Protonen und Neutronen entstehen.

Fragt man nach der experimentellen Prüfung dieser Hypothese der Explosionen, so muß man nach Merkmalen suchen, die gestatten, die Explosionen sicher von den Kaskaden zu unterscheiden. Ein wichtiges Merkmal besteht zunächst darin, daß die Explosion sehr häufig mit einer Kernverdampfung gekoppelt sein wird[1]. Denn beim Zusammenstoß eines sehr energiereichen Teilchens mit einem ruhenden Elektron wird die im Schwerpunktsystem verfügbare Energie stets viel kleiner, als wenn das gleiche Teilchen mit einer ruhenden schweren Partikel zusammenstößt. Beim Zusammenstoß des energiereichen Teilchens mit einem Elektron wird daher die Energie im Schwerpunktsystem im allgemeinen nicht zur Bildung einer größeren Explosion ausreichen, wohl aber beim Zusammenstoß mit einem schweren Teilchen. Da diese meist in Kernen gebunden sind, wird nach der Explosion wahrscheinlich eine Verdampfung des durch die Explosion erwärmten Kerns stattfinden. Als weiteres Merkmal einer Explosion kann gelten, daß viele dabei entstehende Teilchen (wegen ihrer größeren Ruhmasse) keine Kaskaden mehr bilden können. Schließlich bleibt das wichtigste Merkmal der Explosion ihr Auftreten in einer sehr dünnen Schicht. Einige Wilsonaufnahmen von Fussell[2] stellen mit großer Wahrscheinlichkeit kleinere Explosionen dar. Auch zeigt eine eingehende Analyse der Experimente über Hoffmannsche Stöße durch Euler[3], daß in diesen Stößen wahrscheinlich Explosionen eine erhebliche Rolle spielen. Doch müssen hier weitere Experimente abgewartet werden.

Wenn die Explosionen tatsächlich existieren und die für die Konstante r_0 eigentlich charakteristischen Prozesse darstellen, so vermitteln sie vielleicht ein erstes, noch unklares Verständnis der unanschaulichen Züge, die mit der Konstanten r_0 verbunden sind[4]. Diese sollten sich ja wohl zunächst darin äußern, daß die Messung einer Länge mit einer den Wert r_0 unterschreitenden Genauigkeit zu Schwierigkeiten führt. In der Quantentheorie war es die Existenz der Materiewellen oder richtiger das Nebeneinander von Wellen- und korpuskularen Eigenschaften, das dafür sorgte, daß die durch die Unbestimmtheitsrelationen gesetzten Grenzen nicht überschritten werden. In ähnlicher Weise würden die Explosionen dafür sorgen können, daß Ortsmessungen mit einer r_0 unterschreitenden Genauigkeit unmöglich sind. Denkt man z. B. an die Ortsmessung durch ein γ-Strahlmikroskop, so müßten zur Erreichung der gewünschten

---

1) Vgl. hierzu auch W. Heisenberg, Ber. d. Sächs. Ak. d. Wiss. **89**. S. 369. 1938. Insbesondere § 6.

2) L. Fussell, Phys. Rev. **51**. S. 1005. 1937. Für die Übersendung einiger solcher Aufnahmen bin ich Herrn Fussell zu großem Dank verpflichtet.

3) H. Euler, im Erscheinen.

4) Über diesen Zusammenhang verdanke ich Herrn N. Bohr viele lehrreiche Diskussionen.

Genauigkeit γ-Strahlen einer Wellenlänge kleiner als r_0, also Lichtquanten einer Energie größer als $\hbar c / r_0$ verwendet werden. Diese Lichtquanten würden aber an dem zu beobachtenden Gegenstand, auch wenn er eine hinreichend große Masse besitzt — er kann sich dabei in Ruhe oder in Bewegung befinden — im allgemeinen nicht gestreut werden, sondern Explosionen bilden, bei denen die einzelnen entstehenden Teilchen eine Wellenlänge der Größenordnung r_0 besitzen. Eine Abbildung des Gegenstandes mit einer r_0 unterschreitenden Genauigkeit kann dann nicht zustande kommen.

In den letzten Jahren sind zwei verschiedene Versuche unternommen worden, eine universelle Länge r_0 in die Grundlagen des Formalismus der Atomphysik so einzubauen, daß die Divergenzschwierigkeiten der bisherigen Theorien vermieden werden. In mehreren Arbeiten haben Born und Infeld[1]) versucht, die Maxwellsche Theorie im Gebiet kleiner Wellenlängen so abzuändern, daß die Selbstenergie des Elektrons einen endlichen Wert annimmt. Diese Untersuchungen stellen zwar in gewissem Sinne die genaue Erfüllung des Programms der Lorentzschen Elektronentheorie dar, da sie die endliche Ruhmasse des Elektrons in einer relativistisch invarianten und konsequenten Weise berücksichtigen. Sie haben sich jedoch bisher nicht zu einer Quantentheorie des elektromagnetischen Feldes erweitern lassen. Auch nehmen sie wohl zu wenig Rücksicht auf den Umstand, daß die bei der Länge r_0 auftretenden neuen Erscheinungen nicht in der Elektrodynamik, sondern in der Kernphysik ihre Wurzel zu haben scheinen. Ein ganz anderer Versuch zur Beseitigung der Divergenzen ist von March[2]) unternommen worden, der eine Abänderung der Geometrie bei kleinen Längen vorschlägt. Nun entspricht zwar eine solche Abänderung der Geometrie der Vermutung, daß unsere anschaulichen Begriffe nur bis zu Längen der Ordnung r_0 anwendbar sind. Aber es ist die Frage, ob nicht in einem Formalismus, wie dem Marchschen, immer noch zu viele Begriffe der bisherigen Physik unbedenklich verwendet werden; auch ist der Anschluß der Marchschen Vorstellungen an die Erfahrungen der Kernphysik und der Höhenstrahlung bisher nicht erreicht worden.

Wenn man an die umfassenden Änderungen denkt, welche die formale Darstellung der Naturgesetze beim Verständnis der Konstanten c und $\hbar$ erfahren hat, so wird man damit rechnen, daß auch

1) M. Born, Proc. Roy. Soc. (A) **143**. S. 410. 1933; M. Born u. L. Infeld, ebenda **144**. S. 425. 1934; **147**. S. 522. 1934; **150**. S. 141. 1935.

2) A. March, Ztschr. f. Phys. **104**. S. 93 u. 161. 1936; **105**. S. 620. 1937; **106**. S. 49. 1937; **108**. S. 128. 1937.

die Länge r_0 zu völlig neuen Begriffsbildungen zwingt, die weder in der Quantentheorie noch in der Relativitätstheorie ein Analogon besitzen. Insbesondere ist es denkbar, daß es auch hier eine mit Hilfe der Konstanten r_0 formulierbare Invarianzforderung gibt, der alle Naturgesetze zu genügen haben. Vielleicht wird man sich bei dem Versuch, diesen neuen Begriffsbildungen nachzuspüren, zunächst wieder mit Vorteil der Tatsache erinnern, daß es sich in der theoretischen Physik stets nur um die mathematische Verknüpfung beobachtbarer Größen handeln kann; daß uns also einstweilen nur die Aufgabe gestellt ist, Rechenregeln zu finden, durch die wir die Wirkungsquerschnitte der Höhenstrahlungsprozesse teils untereinander, teils mit anderen einfachen Beobachtungsdaten verknüpfen können. Aber zur erfolgreichen Durchführung eines solchen Programms wäre wohl auch eine erhebliche Erweiterung des bisherigen Beobachtungsmaterials die notwendige Voraussetzung.

Leipzig O 27, Bozener Weg 14.

(Eingegangen 13. Januar 1938)

Chapter 27
Hartland S. Snyder (1946): Quantized Space-Time

Hartland S. Snyder (1946). Quantized Space-Time. *Physical Review*, 71: 38–41.

four ⟨111⟩ directions. Pairs of solute atoms lying along these four directions will be affected equally by a shear stress $\sigma_{xx}-\sigma_{yy}$ across the (110) planes, but unequally by a shear stress σ_{xy} across a (001) plane. The first type of shear stress will therefore not cause a preferred orientation of pair axes, i.e., δ' is identically zero, and therefore the Δ_E due to interaction of nearest neighbors is zero when a tensile stress is applied along one of the principal axes. On the other hand, the shear stress σ_{xy} will tend to cause a preferred orientation, and hence δ is not zero. These conclusions may be arrived at in a more elegant manner. Let n_{111}, $n_{\bar{1}11}\cdots$ be the number of pairs of solute atoms with axes lying along the directions [111], [$\bar{1}$11], $\cdots$. The potential energy of the lattice can then contain three interaction terms which are linear in the n's and in the strains. One interaction term will contain the product of the sum of the n's and of the sum $\epsilon_{xx}+\epsilon_{yy}+\epsilon_{zz}$. This term will cause no relaxation, and hence will not be further considered. The remaining two terms represent interaction of the orientation variables with shear strains. These terms must have the same symmetry as the lattice. One term is

$$\alpha\{n_{111}(\epsilon_{yz}+\epsilon_{zx}+\epsilon_{xy})+n_{\bar{1}11}(\epsilon_{yz}-\epsilon_{zx}-\epsilon_{xy})+n_{1\bar{1}1}(-\epsilon_{yz}+\epsilon_{zx}-\epsilon_{xy})+n_{11\bar{1}}(-\epsilon_{yz}-\epsilon_{zx}+\epsilon_{xy})\}.$$

No interaction term can be formed of the ϵ_{xx}, ϵ_{yy}, ϵ_{zz} strains, other than that representing a uniform dilation, which has the cubic symmetry of a cubic lattice.

In f.c.c. lattices the axes passing through nearest neighbors lie along one of the six ⟨110⟩ directions. These pairs are affected unequally by both types of shearing stress, and hence both δ and δ' are different from zero. This conclusion is vindicated by the existence of two shear interaction terms which satisfy the symmetry relations. These are

$$\beta\{(n_{011}+n_{0\bar{1}1})(2\epsilon_{xx}-\epsilon_{yy}-\epsilon_{zz})+(n_{101}+n_{\bar{1}01})(2\epsilon_{yy}-\epsilon_{zz}-\epsilon_{xx})+(n_{110}+n_{\bar{1}10})(2\epsilon_{zz}-\epsilon_{xx}-\epsilon_{yy})\}$$

and

$$\gamma\{(n_{011}-n_{0\bar{1}1})\epsilon_{yz}+(n_{101}-n_{\bar{1}01})\epsilon_{zx}+(n_{110}-n_{\bar{1}10})\epsilon_{xy}\}.$$

PHYSICAL REVIEW VOLUME 71, NUMBER 1 JANUARY 1, 1947

Quantized Space-Time

HARTLAND S. SNYDER
Department of Physics, Northwestern University, Evanston, Illinois

(Received May 13, 1946)

It is usually assumed that space-time is a continuum. This assumption is not required by Lorentz invariance. In this paper we give an example of a Lorentz invariant discrete space-time.

THE problem of the interaction of matter and fields has not been satisfactorily solved to this date. The root of the trouble in present field theories seems to lie in the assumption of point interactions between matter and fields. On the other hand, no relativistically invariant Hamiltonian theory is known for any form of interaction other than point interactions.

Even for the case of point interactions the relativistic invariance is of a formal nature only, as the equations for quantized interacting fields have no solutions. The uses of source functions, or of a cut-off in momentum space to replace infinity by a finite number are distasteful arbitrary procedures, and neither process has yet been formulated in a relativistically invariant manner. It may not be possible to do this.

It is possible that the usual four-dimensional continuous space-time does not provide a suitable framework within which interacting matter and fields can be described. I have chosen the idea that a modification of the ordinary concept of space-time may be necessary because the "elementary" particles have fixed masses and associated Compton wave-lengths.

The special theory of relativity may be based on the invariance of the indefinite quadratic form

$$S^2=c^2t^2-x^2-y^2-z^2, \tag{1}$$

for transformations from one inertial frame to another. It is usually assumed that the variables x, y, z, and t take on a continuum of values and that they may take on these values simultaneously. This last assumption we change to the following:

x, y, z, and t are Hermitian operators for the space-time coordinates of a particular Lorentz frame; the spectrum of each of the operators x, y, z, and t is composed of the possible results of a measurement of the corresponding quantity; the operators x, y, z, and t shall be such that the spectra of the operators x', y', z', t' formed by taking linear combinations of x, y, z, and t, which leave the quadratic form (1) invariant, shall be the same as the spectra of x, y, z, and t.

In other words, we assume that the spectra of the space-time coordinate operators are invariant under Lorentz transformations. It is evident that the usual space-time continuum satisfies the above definition; it is, however, not the only solution. The principal result in this paper is that there exists a Lorentz invariant space-time in which there is a natural unit of length. We hope that the introduction of such a unit of length will remove many of the divergence troubles of present field theory.

It can be shown easily that the introduction of a smallest unit of length in space-time will force one to drop the usual assumptions of commutativity of x, y, z, and t, otherwise the assumption of Lorentz invariance of the spectra of the operators x, y, z, and t, if they commute, implies continuous spectra.

To find operators x, y, z, and t possessing Lorentz invariant spectra, we consider the homogeneous quadratic form[1]

$$-\eta^2=\eta_0^2-\eta_1^2-\eta_2^2-\eta_3^2-\eta_4^2, \tag{2}$$

in which the η's are assumed to be real variables. The variables η_0 to η_4 may be regarded as the homogeneous (projective) coordinates of a real four-dimensional space of constant curvature[2] (a De Sitter space). We now define x, y, z, and t by means of the infinitesimal elements of the group under which quadratic form (2) is invariant. We take

$$\begin{aligned}
x&=ia(\eta_4\partial/\partial\eta_1-\eta_1\partial/\partial\eta_4),\\
y&=ia(\eta_4\partial/\partial\eta_2-\eta_2\partial/\partial\eta_4),\\
z&=ia(\eta_4\partial/\partial\eta_3-\eta_3\partial/\partial\eta_4),\\
t&=(ia/c)(\eta_4\partial/\partial\eta_0+\eta_0\partial/\partial\eta_4),
\end{aligned} \tag{3}$$

in which a is the natural unit of length, c is the velocity of light. The operators x, y, z, and t are assumed to be Hermitian and may be regarded as operating on single valued functions of $\eta_0, \eta_1 \cdots$. From the assumption that x, y, and z are Hermitian operators of the form (3), one can show that each of them has a spectrum consisting of the characteristic values, ma where m is an integer, positive, negative, or zero. The operator, t, has a continuous spectrum extending from minus infinity to plus infinity. The spectrum of each of the operators x, y, z, and t is infinitely degenerate.

Transformations which leave quadratic form (2) and η_4 invariant are covariant Lorentz transformations on the variables η_1, η_2, η_3, and η_0. When the transformed variables η_0', η_1', η_2', η_3', η_4 are substituted into Eqs. (3), it is found that x, y, z, and t undergo contravariant Lorentz transformation. It is evident that the new operators x', y', z', and t' which are formed by replacing η_0, η_1, η_2, η_3, η_4 in Eqs. (3) by η_0', η_1', η_2', η_3', η_4 are linear expressions with real constant coefficients in x, y, z, and t, and are consequently Hermitian operators if x, y, z, and t are Hermitian. The functions on which these operators operate are left invariant except for change in argument. Thus, we see that x', y', z', and t' which are Hermitian operators of the same form in the variables η_0', η_1', η_2', η_3', η_4 as x, y, z, and t were in the variables η_0, η_1, η_2, η_3, η_4 have the same spectra as do x, y, z, t.

We define six additional operators by the equations

$$\begin{aligned}
L_x&=ih(\eta_3\partial/\partial\eta_2-\eta_2\partial/\partial\eta_3),\\
L_y&=ih(\eta_1\partial/\partial\eta_3-\eta_3\partial/\partial\eta_1),\\
L_z&=ih(\eta_2\partial/\partial\eta_1-\eta_1\partial/\partial\eta_3),\\
M_x&=ih(\eta_0\partial/\partial\eta_1+\eta_1\partial/\partial\eta_0),\\
M_y&=ih(\eta_0\partial/\partial\eta_2+\eta_2\partial/\partial\eta_0),\\
M_z&=ih(\eta_0\partial/\partial\eta_3+\eta_3\partial/\partial\eta_0),
\end{aligned} \tag{4}$$

[1] We could also have taken the quadratic form
$$\eta^2=\eta_0^2-\eta_1^2-\eta_2^2-\eta_3^2+\eta_4^2.$$
This leads to another Lorentz invariant discrete space-time.

[2] I am indebted to W. Pauli in a private communication for this interpretation of the variables $\eta_0, \eta_1 \cdots$.

in which h is Plank's constant divided by 2π. Now, L_x, L_y, L_z, M_x, M_y, M_z are, with the exception of the factor ih, the infinitesimal elements of the four-dimensional Lorentz group. It may be shown that each of the above operators commutes with quadratic form (1) when the values of x, y, z, t given by (3) are used. This is another way of stating the Lorentz invariance of quadratic form (1). It will be observed that L_x, L_y, L_z, M_x, M_y, M_z do not involve η_4, and that as a consequence the Lorentz transformation which leaves form (2) and η_4 invariant leaves quadratic form (1) invariant, and as has already been shown, leaves the spectra of the operators x, y, z, and t invariant. We see from these facts that the usual assumptions concerning the continuous nature of space-time are not necessary for Lorentz invariance. This result is the minimum objective of this work.

The ten operators defined in (3) and (4) have a total of forty-five commutators. Only six of these commutators differ from the ordinary ones and these six are

$$
\begin{aligned}
[x, y] &= (ia^2/h)L_z, & [t, x] &= (ia^2/hc)M_x, \\
[y, z] &= (ia^2/h)L_x, & [t, y] &= (ia^2/hc)M_y, \\
[z, x] &= (ia^2/h)L_y, & [t, z] &= (ia^2/hc)M_z.
\end{aligned} \tag{5}
$$

We see from these commutators that if we take the limit $a \rightarrow 0$ keeping h and c fixed, our quantized space-time changes to the ordinary continuous space-time.

The commutators for the quantities L_x, L_y, L_z, M_x, M_y, M_z are those of the infinitesimal elements of the Lorentz group, and for this reason we introduced the factor of ih in their definition. Thus, L_x, L_y, L_z have the usual properties of quantum angular momentum. The commutators which involve one of the operators x, y, z, or t, with one of the operators L_x, L_y, L_z, M_x, M_y, M_z are independent of a, the natural unit of length. Thus all of the commutators pass in the limit $a \rightarrow 0$ to their usual expressions.

In addition to the ten quantities defined in (3) and (4) we define four additional operators, having the transformation properties of space or time displacement operators, or the equivalent, of energy and momentum operators.[3]

$$
\begin{aligned}
p_x &= (h/a)(\eta_1/\eta_4), & p_z &= (h/a)(\eta_3/\eta_4), \\
p_y &= (h/a)(\eta_2/\eta_4), & p_t &= (hc/a)(\eta_0/\eta_4).
\end{aligned} \tag{6}
$$

By means of algebraic manipulations one can show that

$$
L_x = yp_z - zp_y; \quad M_x = \frac{1}{c}xp_t + ctp_x; \text{ etc.} \tag{7}
$$

The angular momenta have their usual expression in terms of coordinate and momentum. One must exercise care as to the order in which the factors in (7) are written since they do not commute. The four quantities p_x, p_y, p_z, p_t commute with one another, and have the same commutators with L_x, L_y, L_z, M_x, M_y, M_z as do the usual expressions for the space or time displacement operators. In addition, each has a continuous spectrum running from minus infinity to plus infinity. Their commutators with the coordinates and time are not the same as usual and are given by

$$
\begin{aligned}
&[x, p_x] = ih[1 + (a/h)^2 px^2]; \\
&[t, p_t] = ih[1 - (a/hc)^2 p_t^2]; \\
&[x, p_y] = [y, p_x] = ih(a/h)^2 p_x p_y; \\
&[x, p_t] = c^2[p_x, t] = ih(a/h)^2 p_x p_t; \text{ etc.}
\end{aligned} \tag{8}
$$

Here we note that if all the components of the momentum are small compared to h/a and the energy is small compared to hc/a then these commutators approach those which are given in orinary quantum mechanics. Further, as we take the limit $a \rightarrow 0$ these commutators change to their standard values.

The fact that these new commutators between coordinate and momentum differ from the old ones appreciably only for large values of the momentum, and that they differ by large amounts when $|p| > h/a$ implies that a field theory based on quantized space-time will give substantially the same results as the usual field

[3] The most general form for the energy momentum operators with the correct transformation properties is

$$
p_x = \frac{h}{a}\frac{\eta_1}{\eta_4} f\left(\frac{\eta_4}{\eta}\right), \cdots, p_t = \frac{hc}{a}\frac{\eta_0}{\eta_4} f\left(\frac{\eta_4}{\eta}\right),
$$

in which $f(\eta_4/\eta)$ is a dimensionless function of its argument. A choice of $f(\eta_4/\eta)$ other than the choice which gives Eqs. (6) may be more useful for physical problems.

theory for all processes which do not involve large components of the momenta, but will produce large effects for processes which do involve large components of the momenta. We might expect that the usual atomic and molecular formulas will suffer no appreciable change, while expressions for self-energy, polarization of the vacuum, and possibly nuclear forces will be considerably altered. Alterations in the last mentioned quantities are certainly necessary.

We note here that the commutator relations (7) have a solution for x, y, z, t.

$$
\begin{aligned}
x &= i\hbar\left[\frac{\partial}{\partial p_x}+(a/\hbar)^2 p_x\left(p_x\frac{\partial}{\partial p_x}+p_y\frac{\partial}{\partial p_y}+p_z\frac{\partial}{\partial p_z}+p_t\frac{\partial}{\partial p_t}\right)\right];\\
t &= i\hbar\left[\frac{\partial}{\partial p_t}-(a/\hbar c)^2 p_t\left(p_x\frac{\partial}{\partial p_x}+p_y\frac{\partial}{\partial p_y}+p_z\frac{\partial}{\partial p_z}+p_t\frac{\partial}{\partial p_t}\right)\right];\ \text{etc.}
\end{aligned}
\tag{9}
$$

This solution of the commutator relations (8) is completely analogous to the solution of the conventional commutator relations between coordinate and momentum for the coordinates in terms of derivatives with respect to the momentum. It would not be surprising to find in our case that no corresponding solution for the momenta in terms of derivatives with respect to the coordinates exists since our coordinates and time do not have continuous spectra in general and do not commute.

It is not difficult to verify that if the expressions for x, y, z, and t given by (9) are substituted in the commutators (5) and the corresponding L_x, L_y, L_z, M_x, M_y, M_z computed, then these last mentioned quantities satisfy their usual commutation relations.

It might appear that x, y, z, and t as given by Eqs. (9) are not Hermitian operators. It may be shown, however, that they are Hermitian operators when the correct volume elements of group space are used. The group which concerns this is the transformation group which leaves quadratic form (2) invariant. An infinitesimal volume element, $d\tau$, of group space is given in terms of p_x, p_y, p_z and p_t by the formula

$$
d\tau=\frac{\hbar dp_x dp_y dp_z dp_t}{ac(p_x^2+p_y^2+p_z^2+(\hbar/a)^2-(p_t/c)^2)^{5/2}}. \tag{10}
$$

To this point the invariance properties of quantized space-time have been considered only with respect to rotations and Lorentz transformations which leave the origin fixed. A continuum of translations is not admissable in this space, indeed one can prove that if x, y, z, and t are such that quadratic form (1) is invariant under infinitesimal displacements, the space x, y, z, and t must be a continuum. Translations of the origin may be introduced as follows. Let $S(\eta_0, \eta_1, \eta_2, \eta_3, \eta_4)$ be an arbitrary unimodular single-valued homogeneous function of the degree zero in the variables, η_0, η_1, η_2, η_3, η_4. Let A be any one of the operators defined by Eqs. (3), (4), or (6). Form the new operators $A'=\bar{S}AS$ in which $\bar{S}$ means the complex conjugate of S. It may then be shown that the primed operators satisfy exactly the same commutation relations as did the original operators and that they are Hermitian. The primed operators thus have exactly the same transformation properties with respect to Lorentz transformations as did the original operators. In particular, the displacement operators p_x, p_y, p_z, p_t are not changed by this process. This is a result which we must demand for the space and time displacement operators. If we choose $S=\exp[im\arctan(\eta_1/\eta_4)]$, we find that this results in a translation of ma in the x direction. It should not be expected, nor is it possible in this space, to find a function S which corresponds to sharply defined translational values of x, y, z, and t simultaneously. The relation between two different frames of reference cannot be set up more precisely than the commutation relations between the coordinates permit.

We will not discuss in any detail in this paper the limitations placed upon the simultaneous measurability of x, y, z, and t due to the noncommutativity of these quantities. Some preliminary calculations which I have made indicate that these limitations are not serious enough to interfere with the ordinary description of atomic phenomena in terms of a continuous space-time nor with our usual macroscopic theory.

Chapter 28
Chen-Ning Yang (1947): On Quantized Space-Time

Chen-Ning Yang (1947). On Quantized Space-Time. *Physical Review*, 72: 874.

Another run was made in position B, where the energy was 88 Mev. This provides the points in Fig. 2 from 88 to 65 Mev. Finally, a run in position A was made, using $2'' \times 3'' \times \frac{1}{4}''$ carbon plates; here the neutron background was relatively large (about 40 percent) because the number of protons was small, but the accuracy was sufficient to show that the cross section does not vary appreciably in going from 88 up to 140 Mev. This last run was not included in the plot.

The striking feature of this result is that there is little if any variation in the $C^{12}(p, pn)C^{11}$ cross section between 60 and 140 Mev. This is not consistent with the picture of proton capture followed by pn evaporation, because the competition of other reactions at high excitation energies would reduce the cross section rapidly. One must assume a non-capture excitation or a (p, n) exchange with roughly constant energy transfer, followed by evaporation of a single neutron or proton. Dr. Serber has pointed out[4] that this is a reasonable assumption in this energy range.

The activity in the boric acid is itself of some interest. The (p, n) reaction in B^{11} gives a peak near the end of the range, as expected for that type of reaction; the second rise at higher energy is attributed to reactions forming C^{11} from oxygen, since its magnitude is consistent with the activity observed in BeO traversed by the same proton current.

This work was done under the auspices of the Atomic Energy Commission, under Contract No. W-7405-Eng-48.

[1] A. C. Helmholz, E. M. McMillan, and D. Sewell, Phys. Rev., to be published.
[2] R. Serber, Phys. Rev., to be published.
[3] W. W. Chupp, E. Gardner, and T. B. Taylor, to be published.
[4] R. Serber, to be published.

On Quantized Space-Time

C. N. Yang
Department of Physics, University of Chicago, Chicago, Illinois
September 15, 1947

RECENTLY Snyder[1] has developed a theory of quantized space-time that is invariant under Lorentz transformations. The theory is, however, not invariant under translations. Indeed, as he has pointed out, to make the theory invariant under translations one must have space-time coordinates forming a continuum.

It does seem desirable, however, to have an invariant theory under a wider group of transformations than the Lorentz transformations. (E.g., we may want to have homogeneity as well as isotropy in space-time.) This can be accomplished by proposing that space-time is curved.

As an example let us consider a de Sitter universe

$$-x_0^2+x_1^2+x_2^2+x_3^2+\xi^2=R^2, \tag{1}$$

which is a pseudosphere in a five-dimensional flat space: x_0, x_1, x_2, x_3, ξ. The group of linear transformations that leaves the quadratic form $-x_0^2+x_1^2+x_2^2+x_3^2+\xi^2$ unchanged could be approximated by the product of the ordinary Lorentz group and the group of translations, for points on the pseudosphere in the region $|x_0| \ll R$, $|x_i| \ll R$. If we define $L_i/\hbar$, $M_i/\hbar$, and $Rp_0/\hbar$ and $Rp_i/\hbar$ as the nuclei of these transformations, we would get the angular- and linear-momenta operators in the de Sitter universe. The p's are given by

$$\begin{aligned} p_0 &= i\hbar\frac{\xi}{R}\frac{\partial}{\partial x_0}+i\hbar\frac{x_0}{R}\frac{\partial}{\partial \xi}, \\ p_i &= -i\hbar\frac{\xi}{R}\frac{\partial}{\partial x_i}+i\hbar\frac{x_i}{R}\frac{\partial}{\partial \xi}. \end{aligned} \tag{2}$$

They satisfy

$$L_i=\xi(x_j p_k - x_k p_j)/R. \tag{3}$$

The commutators involving the p's are

$$[p_i, p_j]=i\hbar L_k/R^2,$$

and

$$[x_i, p_j]=i\hbar\xi\delta_{ij}/R. \tag{4}$$

Those involving the L_i, M_i, and x_μ's are as usual.

The above is, however, not the only theory that is invariant under the Lorentz transformations in five dimensions. We shall propose, in general, that x_0, x_1, x_2, x_3, ξ be Hermitian operators, and that after a Lorentz transformation in five dimensions they can be brought back to their original forms by a unitary transformation. This means that the nuclei of the unitary transformations, which we define to be $L_i/\hbar$, $M_i/\hbar$, and $Rp_0/\hbar$, and $Rp_i/\hbar$, still satisfy the same commutation relations (4). The commutation relations between the operators x_0, x_1, x_2, x_3, ξ are not fixed by the requirement of invariance. If we propose that they are given as follows,

$$\begin{aligned} [x_i, x_j]&=ia^2L_k/\hbar, \quad [x_0, x_i]=ia^2M_i/\hbar, \\ [\xi, x_i]&=ia^2Rp_i/\hbar, \quad [\xi, x_0]=ia^2Rp_0/\hbar, \end{aligned} \tag{5}$$

it is evident that the 15 operators $L_i/\hbar$, $M_i/\hbar$, $Rp_\mu/\hbar$, x_μ/a and ξ/a satisfy the same commutation relations as the nuclei of the group of Lorentz transformations in six dimensions with the basic quadratic form $-\eta_0^2+\eta_1^2+\eta_2^2+\eta_3^2+\eta^2+\zeta^2$. So a possible solution is

$$\begin{aligned} L_i &= i\hbar\left(\eta_k\frac{\partial}{\partial\eta_j}-\eta_j\frac{\partial}{\partial\eta_k}\right), \\ M_i &= i\hbar\left(\eta_0\frac{\partial}{\partial\eta_i}+\eta_i\frac{\partial}{\partial\eta_0}\right), \\ p_0 &= \frac{i\hbar}{R}\left(\eta\frac{\partial}{\partial\eta_0}+\eta_0\frac{\partial}{\partial\eta}\right), \quad p_i = \frac{i\hbar}{R}\left(\eta\frac{\partial}{\partial\eta_i}-\eta_i\frac{\partial}{\partial\eta}\right), \\ x_0 &= ia\left(\zeta\frac{\partial}{\partial\eta_0}+\eta_0\frac{\partial}{\partial\zeta}\right), \quad x_i = ia\left(\zeta\frac{\partial}{\partial\eta_i}-\eta_i\frac{\partial}{\partial\zeta}\right), \end{aligned} \tag{6}$$

and

$$\xi = ia\left(-\zeta\frac{\partial}{\partial\eta}+\eta\frac{\partial}{\partial\zeta}\right).$$

For this solution the eigenvalues of the space coordinates are discrete.

Returning to the general case, we should add Eq. (1) to the general Eqs. (4) and (5).

If we put $R=\infty$ we would get the special solution (2).

[1] H. Snyder, Phys. Rev. 71, 38 (1947).

Chapter 29
Nathan Rosen (1947): Statistical Geometry and Fundamental Particles

Nathan Rosen (1947). Statistical Geometry and Fundamental Particles. *Physical Review*, 72: 298–303.

PHYSICAL REVIEW VOLUME 72, NUMBER 4 AUGUST 15, 1947

Statistical Geometry and Fundamental Particles

NATHAN ROSEN
University of North Carolina, Chapel Hill, North Carolina

(Received May 10, 1947)

To take account of the fact that the existence of a fundamental length a, the classical electron radius, sets a limit to the accuracy with which the position of a point can be measured, it is proposed to introduce two spaces, an "abstract" space consisting of points, and an "observable" space in which one deals with elementary volumes correlated to the points of the former by means of a statistical distribution function in the form of a three-dimensional Gaussian error function. Such a function is not Lorentz invariant, but one can obtain Lorentz covariance in the observable space by carrying out the usual Lorentz transformation in the abstract space. If one assumes that the usual equations for wave fields, in which the fundamental particles are regarded as points, are valid in the abstract space, then one can obtain corresponding equations in the observable space, with the particles behaving as if they had finite volumes. The difficulties associated with infinite self-energies and singularities in the interactions between particles, as calculated by the usual perturbation method, disappear, but the difficulty associated with the divergence of the series of successive orders of perturbations remains.

1. GENERAL CONCEPTS

THE constant a, the "classical electron radius" of the order of 10^{-13} cm, is considered to play an important role in limiting the range of validity of the present form of the quantum theory.[1] It has been suggested by various authors[2] that this constant represents the smallest measurable distance, and that it is therefore necessary to alter the ordinary concepts of geometry when one is dealing with regions having dimensions of the order of a. Accordingly, March[2] introduced a "granular" geometry in which the smallest distinguishable element of space is a sphere of radius a. Recently there appeared a very interesting paper by Snyder,[3] in which the coordinates and time are treated as non-commuting operators, the coordinates having eigenvalues which are integral multiples of a.

In the present paper the point of view adopted is somewhat related to that of March. It is assumed that the constant a determines the lower limit to the error in the measurement of the position of a point. If one measures the x-coordinate of a point under the most favorable conditions, because there does not exist an infinitesimally small measuring rod in nature, one will not obtain an exact value, in general. On the basis of the theory of random errors of measurement it is reasonable to expect that repeated measurements will give values distributed about the mean value in a normal or Gaussian distribution. It will be assumed, then, that the probability of getting a value lying between ξ and $\xi+d\xi$ is given by

$$(2\pi)^{-\frac{1}{2}}a^{-1}\exp[-(\xi-x)^2/2a^2]d\xi,$$

where x is the mean value. Extending this idea to the other coordinates, we will assume that, if one measures the position of a point, for example, that of an electron at a certain instant of time, the probability of obtaining values in the ranges $d\xi$, $d\eta$, $d\zeta$ near ξ, η, ζ is given by

$$dW=\chi(\xi-x,\ \eta-y,\ \zeta-z)d\xi d\eta d\zeta, \qquad (1)$$

where

$$\chi(u,v,w)=(2\pi a^2)^{-\frac{3}{2}}\times\exp[-(u^2+v^2+w^2)/2a^2], \qquad (2)$$

so that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\chi(u,v,w)dudvdw=1. \qquad (3)$$

In this way we arrive at the idea of a geometry dealing with small regions, "elementary volumes," instead of points, provided we think of the elementary volume associated with mean values of the coordinates x, y, z, as made up of those points (values of ξ, η, ζ) for which the function $\chi(\xi-x,\ \eta-y,\ \zeta-z)$ has an appreciable value. Such an elementary volume differs from the kind considered by March in that it does not have a definite boundary.

The basic physical principle associated with such a geometry is that only a physical quantity

[1] W. Heisenberg, Zeits. f. Physik **120**, 513 (1943).
[2] A. March, Naturwiss. **26**, 649 (1938), where references are given to papers by himself and other authors.
[3] H. S. Snyder, Phys. Rev. **71**, 38 (1947).

associated with an elementary volume can be observed, while the value at a point is not directly observable. If one wishes to have physical laws in terms of observable quantities, it is therefore necessary to express them in terms of physical quantities which refer to elementary volumes.

When one attempts to formulate laws in this way, taking account of the basic indeterminacy of position, it is convenient to work with two spaces, an "abstract" or point space and an "observable" space, made up of elementary volumes, as described above. To every point of the abstract space one can correlate an elementary volume of the observable space by taking the mean values of the coordinates of the latter equal to the corresponding coordinates of the former. In the observable space, no meaning will be assigned to the value of a physical quantity, such as a field variable, at a point, but only to the mean value over an elementary volume. The purpose of introducing the abstract space is to help one to obtain such mean values.

In attempting to formulate physical laws in the framework of this statistical geometry it is natural at the outset to try to keep them as nearly as possible in the same form as at present. Thus, in the case where one is dealing with physical variables satisfying linear equations, one might try to retain these equations, but to take them as holding in the *abstract space*. The corresponding equations in the observable space would then be obtained by an averaging process from the latter, so that, for example, a function $f(x, y, z, t)$ would go over into[4]

$$\bar{f}(x, y, z, t) \equiv \langle f(x, y, z, t) \rangle$$

according to the relation

$$\bar{f}(x, y, z, t) = \int\int\int f(\xi, \eta, \zeta, t) \times \chi(\xi - x, \eta - y, \zeta - z) d\xi d\eta d\zeta. \quad (4)$$

Such a transformation has some interesting properties. In the case of a constant, C, it follows obviously that

$$\bar{C} = C. \quad (5)$$

One also sees readily that for a coordinate x and the time t

$$\bar{x} = x, \quad \bar{t} = t. \quad (6)$$

In the case of any well-behaved function $f(x, y, z, t)$ it follows that

$$\langle Cf \rangle = C\bar{f}, \quad \langle f_1 + f_2 \rangle = \bar{f}_1 + \bar{f}_2 \quad (7)$$

$$\langle \partial f/\partial x \rangle = \partial \bar{f}/\partial x, \quad \langle \partial f/\partial t \rangle = \partial \bar{f}/\partial t, \quad (8)$$

and

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \bar{f} dx dy dz = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f dx dy dz. \quad (9)$$

On the other hand one finds that

$$\langle xf \rangle = x\bar{f} + a^2(\partial \bar{f}/\partial x). \quad (10)$$

It follows that for a function $f(x, y, z, t)$ which satisfies a linear homogeneous partial differential equation with constant coefficients, the equation for $\bar{f}(x, y, z, t)$ will have the same form as that for f. It should be noted however that, while we have obtained the form of the equation in the observable space by an averaging process, this does not mean necessarily that each solution $\bar{f}$ must be obtained by first determining a suitable solution f and then getting its mean value over an elementary volume. Once the appropriate equation in the observable space has been found, it is possible to get its solutions directly. Each such solution, however, is subject to the restriction that it must be possible to express it as the mean value of some function over the elementary volume.

2. LORENTZ TRANSFORMATION

Before considering applications of the preceding ideas, it is necessary to examine them from the standpoint of the special theory of relativity. If we have two frames of reference moving with a uniform relative velocity, it is to be expected that the same uncertainty in the measurement of position will exist in each of them. In each system one will have "elementary volumes," and an observer in each one will be able to introduce an "abstract" and an "observable" space. Any physical law must be expressible in the same form in each frame of reference.

Now, the special relativity theory was developed without taking account of the limitation on measurement imposed by the existence of the

[4] Owing to typographical difficulties in setting bars over multiple symbols or boldface symbols, the notation with angular brackets will be used where required. *Ed.*

fundamental constant a. Hence one should be prepared to find that the introduction of this constant into the theory leads to changes in the usual transformation relations between quantities in the two frames of reference.

In order to be sure of obtaining transformation equations that will lead to physical laws of the same form in the two systems, one can make the following assumption: The usual Lorentz transformation will be assumed to hold in the transformation from the abstract space of one reference frame to that of the other frame moving relative to it with a uniform velocity. The transformation of a physical quantity in the observable space can then be determined from the corresponding transformation in the abstract space, if one knows how to go from the abstract to the observable space.

It may turn out in some cases that the transformation in the observable space will differ from the Lorentz transformation. However, as far as the transformation of coordinates and time is concerned, we see from Eq. (6) that the Lorentz transformation will remain valid. From a consideration of a plane monochromatic wave one can readily see that the frequency and wavelength of a light wave transform in the usual way.

At this point it is appropriate to consider briefly another question suggested by relativistic considerations. In the preceding discussion the time coordinate has been treated quite differently from the space coordinates. In particular, we have considered the uncertainty in the measurement of a space coordinate at a particular moment of time. It might be asked whether one should not introduce an uncertainty in the time (in this case of the order of a/c) in the same way as has been done for the coordinates. This would mean having on the right-hand side of Eq. (1) an additional factor

$$(2\pi)^{-\frac{1}{2}}ca^{-1}\exp[-c^2(\tau-t)^2/2a^2]d\tau.$$

However, the objection to such a procedure appears to be that it would be difficult to interpret the formalism from the operational standpoint. In the three-dimensional treatment considered previously the function χ of Eq. (1) represents the probability distribution of measured values of the coordinates of a point, such as the position of an electron, at a given moment of time, so that the time serves as a parameter to identify what is to be measured. In the four-dimensional treatment, where the time would also be measured and would have an uncertainty, no such parameter would in general be available, so that it would not always be clear what one was measuring or how the measurement could be repeated. The procedure adopted here, of using the time as a parameter is, after all, in agreement with what is generally done in the quantum theory.

3. ELECTRON AND RADIATION FIELD

We next consider the problem of the electron and its interaction with the electromagnetic field. In order to avoid lengthy derivations, reference will be made to the book of Heitler[5] for a discussion of the classical and quantum theories of the electromagnetic field.

The electromagnetic field for a given distribution of charges with a density ρ and convective velocity $\mathbf{v}$ can be described classically in terms of the potentials ϕ, $\mathbf{A}$ by means of the following equations:[6]

$$\square^2\phi=-4\pi\rho, \tag{11}$$

$$\square^2\mathbf{A}=-(4\pi/c)\rho\mathbf{v}, \tag{12}$$

$$\nabla\cdot\mathbf{A}+\dot{\phi}/c=0, \tag{13}$$

with a dot denoting differentiation with respect to the time. The electric and magnetic field intensities $\mathbf{E}$ and $\mathbf{H}$ are then given by

$$\mathbf{E}=-\nabla\phi-\dot{\mathbf{A}}/c,\quad \mathbf{H}=\nabla\times\mathbf{A}. \tag{14}$$

If we assume that these equations are to remain valid in the abstract space and then go over to the observable space by transformations of the type of Eq. (4), we get in place of (11)

$$\square^2\bar{\phi}=-4\pi\bar{\rho}, \tag{15}$$

while the other equations are changed in the same way. We see that in the observable space the equations for the electromagnetic field have the same form as in the abstract space, except for the fact that the charge and current densities refer to mean values over an elementary volume. Hereafter, we shall drop the bars from the field

[5] W. Heitler, *The Quantum Theory of Radiation*, second edition (Oxford Press, England, 1944).

[6] The d'Alembertian operator $\square^2\equiv\nabla^2-\partial^2/c^2\partial t^2$.

variables, but it will be understood that they refer to the observable space.

It has been suggested[7] that an electron, because of its fundamental character, should be represented by the simplest of all structures, a point. Let us adopt this standpoint, but let us represent the electron by a point in the *abstract space*. The corresponding representation in the observable space, however, will no longer be a point but an elementary volume.

Consider a point charge e_k having at the time t coordinates x_k, y_k, z_k, and an instantaneous velocity $\mathbf{v}_k$ in the x direction. The charge density ρ at a point x, y, z, at the time t is then given by

$$\rho = e_k(1-v_k^2/c^2)^{-\frac{1}{2}}\delta([x-x_k][1-v_k^2/c^2]^{-\frac{1}{2}}) \times\delta(y-y_k)\delta(z-z_k), \quad (16)$$

where δ is the Dirac delta-function. It follows that

$$\bar{\rho} = e_k\chi(x-x_k,\, y-y_k,\, z-z_k), \quad (17)$$

so that the "effective charge density" is now given by a Gaussian distribution. It is seen that $\bar{\rho}$ has spherical symmetry about the center. Thus this effective charge distribution does not undergo a Lorentz contraction when it is in motion. This is, of course, attributable to the fact that the charge distribution arises from the geometry of the space and not from any assumed internal structure of the electron. A similar behavior was pointed out by March.[2]

In the same way one obtains in the observable space

$$\langle\rho\mathbf{v}\rangle = e_k\mathbf{v}_k\chi(x-x_k,\, y-y_k,\, z-z_k), \quad (18)$$

where $\mathbf{v}_k$ is the velocity of the point charge in the abstract space, or that of the center of the charge distribution in the observable space.

We see then that, on the basis of statistical geometry, one obtains a description of the electron in which many of the previous difficulties have been removed. The electron behaves like a distributed charge without any singularities, as far as its interaction with the electromagnetic field is concerned. At the same time any questions concerning the stability of the electron or the nature of the cohesive forces holding the charge together drop out. It is meaningless to ask about the force exerted on or by a part of the electron, but only the electron as a whole.

The equations for the electromagnetic field associated with a system of electrons can now be written

$$\Box^2\phi = -4\pi\sum_k e_k\chi(x-x_k,\, y-y_k,\, z-z_k), \quad (19)$$

$$\Box^2\mathbf{A} = -(4\pi/c)\sum_k e_k\mathbf{v}_k \times\chi(x-x_k,\, y-y_k,\, z-z_k), \quad (20)$$

with the remaining equations having the same form as (13) and (14).

Let now expand ϕ and $\mathbf{A}$ in series of plane monochromatic waves, following Heitler (reference 5, p. 47) and using his notation. For example, if we write

$$\phi = \sum_\sigma a_\sigma(t)\phi_\sigma(x, y, z), \quad (21)$$

then from (19) we obtain

$$\ddot{a}_\sigma + \omega_\sigma^2 a_\sigma = \sum_k e_k\bar{\phi}_\sigma(k), \quad (22)$$

where ω_σ is the (circular) frequency, and $\bar{\phi}_\sigma(k)$ is given by

$$\bar{\phi}_\sigma(k) = \int\phi_\sigma(\xi, \eta, \zeta) \times\chi(\xi-x_k,\, \eta-y_k,\, \zeta-z_k)d\xi d\eta d\zeta, \quad (23)$$

which is the mean value of ϕ_σ over the elementary volume of the k'th particle. Equation (22) is the same as the corresponding equation given by Heitler (reference 5, p. 49) except for the fact that here we have $\bar{\phi}_\sigma$ instead of ϕ_σ on the right-hand side. In the same way, in the expansion of the vector potential $\mathbf{A}$ in series of transverse and longitudinal plane waves, one gets equations for the coefficients which differ from those given there by Heitler only with respect to this averaging over the elementary volume.

Now the effect of averaging over the elementary volume becomes important as one goes to high frequencies. Consider a plane monochromatic wave, say, of the form

$$\phi = B\exp[i(kx-\omega t)], \quad (24)$$

with $\omega = kc$. Then one finds

$$\bar{\phi} = B[\exp(-k^2a^2/2)]\exp[i(kx-\omega t)] = [\exp(-k^2a^2/2)]\phi = [\exp(-\omega^2a^2/2c^2)]\phi. \quad (25)$$

[7] J. Frenkel, Zeits. f. Physik **32**, 518 (1925); P. A. M. Dirac, Proc. Roy. Soc. **A167**, 148 (1938).

This means that the coupling of an electron with the high frequency components of the electromagnetic field is now weaker than in the classical theory.

If again one follows Heitler (reference 5, p. 50) in writing down a Hamiltonian function for the system and then calculating the energy present in the longitudinal field components, one finds that this can be expressed as a sum of terms of two kinds: (1) interaction energies of pairs of particles, and (2) self-energies of single particles. For the energy of interaction of two particles one obtains, on integrating over the various frequencies and directions of the plane waves,

$$V_{ik}=(e_ie_k/r)\Phi(r/2a), \quad (26)$$

where e_i, e_k are the charges of the particles, r is the distance between them, and

$$\Phi(x)=2\pi^{-\frac{1}{2}}\int_0^x[\exp(-u^2)]du. \quad (27)$$

For $r\gg a$, this goes over into the coulomb interaction, but as r approaches zero, it remains finite, instead of going to infinity as does the coulomb interaction.

The self-energy of a particle W_k one obtains either by going through the same kind of calculation as that leading to Eq. (26) or by getting $\frac{1}{2}V_{kk}$ from (26), letting r approach zero. One finds for the self-energy

$$W_k=\tfrac{1}{2}V_{kk}=e_k^2/2\pi^{\frac{1}{2}}a. \quad (28)$$

Incidentally, the same results for V_{ik} and W_k can be obtained by calculating the electrostatic interaction and electrostatic energy of charge distributions of the form given by Eq. (17).

If one equates the expression (28) for an electron to its rest energy m_0c^2 one finds

$$a=0.79\times10^{-13}\ \mathrm{cm}$$

based on the way a was defined in Eq. (2). However, it is not clear at present that such a procedure of equating the electrostatic self-energy to the total rest energy is justified.

Let us now consider the question of quantization. In view of what was found for the equation of the electromagnetic field in the observable space it is plausible to take for the quantum-theoretical equations for the system consisting of electrons and radiation field the same equations used up to now (reference 5, Chapters II and III) with the sole modification that in the interaction of an electron with the field, the mean value of the latter over an elementary volume must now be used.

The Hamiltonian for the system can be written

$$H=\sum_k H_k+\sum_\lambda H_\lambda+\sum_{i>k} V_{ik}+\sum_k W_k+H'. \quad (29)$$

Here the Hamiltonian for a particle is given on the basis of the Dirac theory by

$$H_k=c\alpha_k\cdot(\mathbf{p}_k-e_k\langle\mathbf{A}^e\rangle(k))+m_kc^2\beta_k+e_k\bar{\phi}^e(k), \quad (30)$$

$\bar{\phi}^e(k)$, $\langle\mathbf{A}^e\rangle(k)$ being the mean values of the potentials of the external field at the k'th particle. The Hamiltonian for a transverse radiation wave is given by

$$H_\lambda=2\omega_\lambda^2q_\lambda^*q_\lambda \quad (31)$$

in terms of the quantized amplitudes q_λ (a discussion of which will be found in reference 5, p. 59). The term H' represents the interaction between the electrons and the transverse, or radiation, field. It is given in the present case by

$$H'=-\textstyle\sum_k e_k\boldsymbol{\alpha}_k\cdot\langle\mathbf{A}\rangle(k)$$
$$=-\sum e_k\boldsymbol{\alpha}_k\cdot(q_\lambda\langle\mathbf{A}_\lambda\rangle(k)+q_\lambda^*\langle\mathbf{A}_\lambda^*\rangle(k)), \quad (32)$$

where

$$\langle\mathbf{A}_\lambda\rangle(k)=\int\mathbf{A}_\lambda(\xi,\eta,\zeta)\times\chi(\xi-x_k,\eta-y_k,\zeta-z_k)d\xi d\eta d\zeta, \quad (33)$$

and $\mathbf{A}_\lambda$ is the vector potential for a plane monochromatic transverse wave, suitably normalized (reference 5, p. 59). It follows from Eq. (25) that

$$\langle\mathbf{A}_\lambda\rangle(k)=\exp[-\omega_\lambda^2a^2/2c^2]\mathbf{A}_\lambda(x,y,z). \quad (34)$$

It might be pointed out that in (31) we are taking the Hamiltonian of the radiation field in the usual form. An alternative procedure is based on the supposition that the quantization of the radiation field is carried out in the abstract space instead of the observable space. This, however, necessitates giving up the proportionality between energy and frequency and leads to the existence of a maximum value for the energy of a photon. Such a situation does not appear satisfactory.

If one treats the coupling term H' as a perturbation and carries out the usual perturbation-theory calculations, one finds that the use of mean values of the radiation field over an elementary volume, as given by Eq. (34), gets rid of the previous "high frequency" divergence difficulties.

We have already seen how one obtains a finite value for the static self-energy, Eq. (28). By means of a second-order perturbation calculation one can calculate the dynamic (or transverse) self-energy (reference 5, p. 181). One finds, on taking account of the negative energy states of the electron, that this is given by

$$W_k' = (e_k^2\hbar c/2\pi a^2 E) - (e_k^2/2a\pi^{\frac{1}{2}}) \times[1-(m_0^2c^3/2pE)\ln\{(E+cp)/(E-cp)\}], \quad (35)$$

where E is the unperturbed energy and p the momentum of the electron. This result, while finite, is unsatisfactory, however. The first term, which is the important one for small velocities, turns out to have a value which is much larger than m_0c^2, if one takes $a \sim 10^{-13}$ cm.

It appears that this is related to another difficulty: if one calculates the contribution to the self-energy of higher order perturbation terms, one finds that, while the value obtained for each order is finite, the terms of even order (the only non-vanishing terms) keep getting larger as one goes to higher orders, for $a \sim 10^{-13}$ cm. A rough estimate indicates that the ratio of an even-order term to the preceding one is of the magnitude of $e^2hc/a^2(mc^2)^2$, or of 137 for a equal to the classical electron radius. This means, of course, that the perturbation calculation is divergent and hence should not be used in determining the energy of interaction of the electron and the field. It appears that the coupling between the electron and the field is too strong for the perturbation theory to be applicable. It might be pointed out that this difficulty is not peculiar to the present approach. In a sense a similar difficulty also exists in the usual "point-electron" theory, but is masked by the presence of infinities in the individual terms of the perturbation calculation.

As has already been remarked, in the observable space the interaction between a particle and the field involves the mean value of the field over the elementary volume of the particle. Since our knowledge of the field can be obtained only from the observation of its effect on particles, it follows that only such a mean value (a mean of the mean, from the standpoint of the abstract space) can be determined. Hence it is to be expected that in such matters as commutation relations among field variables, etc., one should deal with such mean values over elementary volumes, and not the values at a point.

4. OTHER PARTICLES

In the case of other elementary particles one can use the same method as for electrons: the particle is regarded as a point in the abstract space and, therefore, as an elementary volume in the observable space. This leads to the expectation that all fundamental particles should have the same "size."

The expression for the interaction of a particle with a field in the observable space must take account of the fact that the particle is now an extended source. This can be done as above, by writing down the field equations in the abstract space and then going over to the observable space. From the form of the equations the corresponding Hamiltonian can then be deduced. In general it differs from the Hamiltonian for "point" particles in that the interaction terms involve the average values of the field variables over the elementary volumes of the particles.

In the case of nucleons interacting with a meson field, it is evident that the static interaction between two heavy particles (calculated in first approximation) will not have any singularity. Hence, if such an interaction is used as the potential energy in the Schrödinger equation, no "cutting off" of this interaction energy at small distances between the particles is necessary.

Note added in proof:—It has been kindly pointed out to me by Mr. M. F. M. Osborne that in the book by A. S. Eddington, *Fundamental Theory* (Cambridge University Press, 1946), use was made of an "abstract" space and a Gaussian-error transformation function, although in somewhat different applications.

Chapter 30
Alfred Schild (1948): Discrete Space-Time and Integral Lorentz Transformations

Alfred Schild (1948). Discrete Space-Time and Integral Lorentz Transformations. *Physical Review*, 73: 414–415.

single element or on a group of elements, such as those comprising the kth plane. Thus

$$w_{k,f}=\phi(f_k), \tag{3}$$

where f_k = component of force in the direction of the observed stress acting over the *whole* kth cross section. Similarly, for a backward unit dislocation,

$$w_{i,k,b}=\psi(f_{i,k}), \tag{4}$$

and

$$w_{k,b}=\psi(f_k). \tag{5}$$

In view of Eqs. (1) to (5) we have

$$\left.\begin{aligned}\phi(f_k)&=\prod_{i=1}^{N_k}\phi(f_{i,k}),\\ &\text{and}\\ \psi(f_k)&=\prod_{i=1}^{N_k}\psi(f_{i,k}).\end{aligned}\right\} \tag{6}$$

But, clearly,

$$f_k=\sum_{i=1}^{N_k} f_{i,k}. \tag{7}$$

In accordance with a proof frequently cited in thermodynamics,[1] Eqs. (6) and (7) define the exponential relations:

$$\begin{aligned}w_{k,f}&=\phi(f_k)=\exp c_1f_k,\\ &\text{and}\\ w_{k,b}&=\psi(f_k)=\exp c_2f_k,\end{aligned} \tag{8}$$

where c_1 and c_2 are parameters.

Assumption II. The probability of a unit dislocation per unit time in the direction of the force component (forward jump) is the same in either direction from the cross-sectional plane, i.e.,

$$\phi(f_k)=\psi(-f_k). \tag{9}$$

Now, since a forward jump and a backward jump are mutually exclusive events, the net probability of a forward jump per unit time is the difference $w_{k,f}-w_{k,b}$. The net strain in the direction of stress, occurring in time Δt, is evidently

$$\Delta\epsilon=n\lambda(w_{k,f}-w_{k,b})\Delta t. \tag{10}$$

From Eqs. (8) and (9) it follows that $c_1=-c_2$, and hence we have from Eqs. (8),

$$\left.\begin{aligned}w_{k,f}-w_{k,b}&=\exp c_1f_k-\exp(-c_1f_k)\\ &=2\sinh c_1f_k.\end{aligned}\right\} \tag{11}$$

Thus,

$$\Delta\epsilon=n\lambda\times 2\sinh c_1f_k\times\Delta t,$$

or

$$\lim_{\Delta t\to 0}\frac{\Delta\epsilon}{\Delta t}=\frac{d\epsilon}{dt}=2n\lambda\sinh c_1f_k. \tag{12}$$

Placing $f_k=a\sigma$, $2n\lambda=A$, and $ac_1=B$, we finally obtain

$$d\epsilon/dt=A\sinh B\sigma, \tag{13}$$

where σ = the stress in the specimen in the neighborhood of the kth plane.

Equation (13) is of the form which has been proposed by several workers[2-4] for the description of plastic deformational phenomena in metals as well as high polymers. While the foregoing theory itself gives no description of the elementary mechanism of dislocation, Assumptions I and II may be regarded as establishing conditions which must be met by postulated molecular forces in a more detailed theory leading to Eq. (13).

[1] Cf., for example, Max Planck, *Theory of Heat* (MacMillan Company, Ltd., London, 1932), p. 227.
[2] A. Nadai, *Stephen Timoshenko Anniversary Volume* (MacMillan Company, Ltd., London, 1938), p. 155.
[3] A. Tobolsky and H. Eyring, J. Chem. Phys. **11**, 125 (1943).
[4] A. V. Tobolsky and R. D. Andrews, J. Chem. Phys. **13**, 3 (1945).

Discrete Space-Time and Integral Lorentz Transformations

A. Schild[1]
The Institute for Advanced Study, Princeton, New Jersey
December 22, 1947

THE idea of introducing discreteness into space and time has occasionally been considered[2] as a means of removing the "infinities" which trouble modern physical theory, both classical and quantal. The objection which is usually raised against such discrete schemes is that they are not invariant under the Lorentz group. The purpose of this investigation is to show that *there is a simple model of discrete space-time which, although not invariant under all Lorentz transformations, does admit a surprisingly large number of Lorentz transformations.* This group of transformations is, in fact, sufficiently large to make conceivable the use of this model as a background for physical theory.

Consider all events in Minkowski space-time whose four coordinates $x^r\equiv(t, x, y, z)$ are integers.[3] (The velocity of light is taken as unity.) We shall call the set of such events the "hypercubic lattice," or *cubic lattice,* for short. The proper homogeneous Lorentz transformations which leave the cubic lattice as a whole invariant will be called *integral Lorentz transformations.* A world vector whose four coordinates are integers will be called an *integral vector*; the adjective *primitive* will be added if the components of the vector have no common integer factor other than ± 1. Two integral vectors obtainable from one another by an integral Lorentz transformation will be called *integral transforms.* We now give a brief sketch of the principal results; a more detailed account will be published elsewhere.

The integral Lorentz transformations form a group. A Lorentz transformation $x'^r=L_s{}^rx_s$ is integral if and only if all its components $L_s{}^r$ are integers. Thus the problem of finding all integral Lorentz transformations is equivalent to the solution of 10 quadratic Diophantine equations in 16 unknown integers $L_s{}^r$. This problem can be approached by means of the two-dimensional spinor calculus[4] and the elementary theory of Gaussian integers.[5]

With an integral vector (t, x, y, z) we associate a Hermitian spintensor $a^{\dot\alpha\beta}$ as follows:

$$\begin{aligned}a^{\dot 1 1}&=t+z, & a^{\dot 1 2}&=x-iy,\\ a^{\dot 2 1}&=x+iy, & a^{\dot 2 2}&=t-z.\end{aligned} \tag{1}$$

It can be shown that all primitive integral null vectors pointing into the future are represented by spinvectors c^α:

$$a^{\dot{\alpha}\beta} = c^{\dot{\alpha}} c^\beta, \tag{2}$$

where c^1, c^2 are Gaussian integers, so that c^1 and c^2 are relatively prime and neither is divisible by $1+i$, or else so that $1+i$ is the greatest common factor of c^1, c^2 and one of them is divisible by 2. Such spinvectors will be called *integral*.

As is well known, the spin transformations

$$c'^\alpha = \lambda_\beta{}^\alpha c^\beta, \quad c'^{\dot{\alpha}} = \bar{\lambda}_\beta{}^\alpha c^{\dot{\beta}}, \quad |det(\lambda_\beta{}^\alpha)| = 1, \tag{3}$$

represent all proper homogeneous Lorentz transformations, $\lambda_\beta{}^\alpha$ being determined by $L_s{}^r$ to within an arbitrary phase factor $e^{i\theta}$. If this phase factor is chosen suitably, the integral Lorentz transformations are represented by exactly those spin transformations which, together with their inverse, map integral spinvectors into integral spinvectors. This theorem enables us to find the following spin representation of the integral Lorentz group:

$$\lambda_1{}^1\lambda_2{}^2 - \lambda_2{}^1\lambda_1{}^2 = 1, \tag{4}$$

where one of the following cases applies: I. $\lambda_\beta{}^\alpha$ are Gaussian integers such that $\Sigma_{\alpha\beta}\lambda_\beta{}^\alpha$ is divisible by $1+i$. II. $\lambda_\beta{}^\alpha = \mu_\beta{}^\alpha/(1+i)$, where $\mu_\beta{}^\alpha$ are Gaussian integers not divisible by $1+i$. III. $\lambda_\beta{}^\alpha = 2^{\frac{1}{2}}\mu_\beta{}^\alpha/(1+i)$, where $\mu_\beta{}^\alpha$ are Gaussian integers such that $\Sigma_{\alpha\beta}\,\mu_\beta{}^\alpha$ is divisible by $1+i$. IV. $\lambda_\beta{}^\alpha = \frac{1}{2}2^{\frac{1}{2}}\mu_\beta{}^\alpha$, where $\mu_\beta{}^\alpha$ are Gaussian integers not divisible by $1+i$. The theory of Gaussian integers shows immediately that each of the above cases includes an infinity of spin transformations. Thus, *the integral Lorentz group is infinite*, though discrete.

From the above it can be deduced that all primitive integral null vectors are equivalent, in the sense that any two of them are integral transforms of one another.

Consider any integral vector and form the set of all its integral transforms. Project each of these vectors onto the xyz space. Then the directions defined by these projections in 3-space are everywhere dense. This shows that our discrete space-time model possesses a large measure of *spatial isotropy*. It is obvious that our cubic lattice is invariant under all translations which map one lattice point into another. In this sense our discrete model is *homogeneous*.

Finally we must mention a property of our model which constitutes a drawback as far as hopes for physical application are concerned. The velocities associated with integral Lorentz transformations are given by the formula $v = (n^2-1)^{\frac{1}{2}}/n$, where n is any positive integer. The smallest non-zero velocity is $\frac{1}{2}3^{\frac{1}{2}} = 0.866$ times the velocity of light.

[1] Frank B. Jewett Fellow, on leave from Carnegie Institute of Technology, Pittsburgh, Pennsylvania.
[2] V. Ambarzumian and D. Iwanenko, Zeits. f. Physik **64**, 563 (1930); L. Silberstein, *Discrete Space-Time* (University of Toronto Studies, Physics Series, 1936).
[3] The coordinates are integral multiples of a "fundamental length" ϵ (probably of the order of nuclear dimensions). We choose ϵ as the unit of length.
[4] O. Laporte and G. E. Uhlenbeck, Phys. Rev. **37**, 1381 (1931).
[5] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers* (Oxford University Press, New York, 1908), Chapter XII.

Range and Energy of Beta-Radiation from Calcium 45

A. K. Solomon
Biophysical Laboratory, Harvard Medical School, Boston, Massachusetts
AND
L. E. Glendenin
Massachusetts Institute of Technology, Cambridge, Massachusetts
December 29, 1947

THE radiations given off by Ca^{45} have been investigated by Walke, Thompson, and Holt[1] who report beta-radiations of maximum energy 0.2 and 0.9 Mev, gamma-radiation of 0.7 Mev, and a half-life of 180 days. The energies of the radiation were determined by absorption measurements. In view of the importance of this isotope in biological research, it has seemed advisable to reinvestigate the radiation characteristics.

Experimental details.—Carrier-free Ca^{45}, produced by $n-p$ reaction on monoisotopic Sc^{45}, was obtained from the Atomic Energy Commission. The counting apparatus was the same as that previously described;[2] a thin-window (1.9 mg/cm²) Geiger counter was used. The source was essentially carrier-free, and was deposited in a thin (0.017-inch) aluminum stamping. The beta-radiation was measured by absorption in aluminum foils.

Results and discussion.—The method of Feather[3] was used in analyzing the results, as previously described.[2] The initial strength of the sources varied from 3000 to

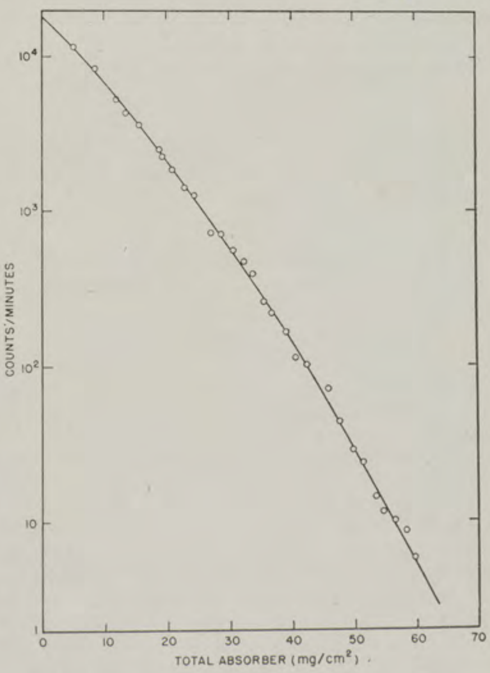

FIG. 1. Aluminum absorption curve of Ca^4 β-radiation.

Gauge and Constraints

Chapter 31
The Genesis of Canonical Quantum Gravity

Alexander Blum and Donald Salisbury

The late 1940s saw the solution of the divergence difficulties of quantum field theory, or at least quantum electrodynamics.[1] The renormalization program was successful not only in removing the infinities, but also in giving precise values to the finite remainders. These were in excellent agreement with newly discovered precision effects, such as the Lamb shift and the anomalous magnetic moment of the electron. The renormalization techniques were, at least initially, based on the covariant quantization procedure discussed in Chapter 17 and their application to the quantization of gravity was taken up in the 1950s.

In parallel, and sometimes in direct opposition, to the development of renormalized quantum field theory (QFT), a second route to quantum gravity was developed in the late 1940s, initially by Peter Bergmann. Bergmann had been an assistant of Einstein from 1936 to 1941 and had worked with Einstein on the attempts to construct a five-dimensional unified field theory (UFT).[2] But unlike Einstein, Bergmann was not in radical opposition to modern quantum theory, having learned it as a student of Philipp Frank in Prague in the mid-1930s. After obtaining a permanent position at Syracuse University, Bergmann thus abandoned the pursuit of a unified field theory and set himself to reconcile general relativity (GR) and quantum theory through the construction of a quantum theory of gravity.

Bergmann was little concerned with the contemporary rapid development of renormalized quantum field theory. And indeed, the covariant QED of the day did not seem to harmonize well with his vision of quantum gravity. The QED of Schwinger was construed as the quantization of free, linear field theories, both for matter and the electromagnetic field, which were then perturbed by the introduction of interactions between the fields, leading to the non-trivial time evolution of the quantum state in the interaction picture.

But for Bergmann the essential element of GR was its non-linearity, because it made possible the derivation of the equations of motion of point particles directly from the field equations, as outlined by Einstein, Infeld and Hoffmann (EIH) during Bergmann's time with Einstein. Bergmann's hope was that this feature would carry over from the classical to the quantum theory, so that a quantum theory of gravity would directly lead to the quantum equations of motion of point particles as well. He was thus envisioning not merely an appropriation of the field theory of general relativity through quantum theory by the methods of field quantization, but rather a true marriage between general relativity and quantum theory, where the introduction of general relativity would help to solve the divergence difficulties of quantum field theory, which he, like many others, saw as originating from the difficulties of having point-like particles.

Of course this approach was not entirely up to date with the current state of quantum field theory, which had in a sense dispensed with the quantum mechanical notion of in principle localizable, point-like entities. Photons were certainly not perfectly localizable, nor were charged particles in electron-positron theory. And Bergmann had to admit that

[1] See Schweber (1994).
[2] On Bergmann's background, see Halpern (2005).

already classically the EIH procedure was only feasible if one disregarded splitting world lines, as in the emission of a photon or electron-positron pair creation. But these difficulties were put aside for the moment. Bergmann was after a quantum theory of point particles and the gravitational field.

The predilection for the EIH procedure was not the only inheritance from Bergmann's time with Einstein. From his work on UFT, he also had a clear understanding that GR was not to be considered the final field theory of gravitation. In his program, he thus did not engage directly with GR and the quantization of the space-time metric, but preferred to discuss the quantization of a more general field theory that kept what he considered to be the essential elements of GR: general covariance and non-linearity, which together ensured the feasibility of the EIH-type determination of particle trajectories.

The study of the quantization of a field theory mainly characterized by its invariance properties unwittingly brought him very close to Rosenfeld's first 1930 paper on what we have called the momentum difficulty. And indeed, in his first paper on the subject (Chapter 32), Bergmann showed not only that a general non-linear, generally covariant field theory would have all the attractive features of GR (EIH determination of the equations of motion, as well as energy conservation, at least in the same sense as in GR), but would also lead to identities involving the canonical momenta, of the type that Rosenfeld had first discovered. There are slight differences in the treatment, of course: Bergmann did not consider spinorial matter and thus did not introduce tetrads or local Lorentz symmetry.[3] Without tetrads, even for the case of regular general relativity, Bergmann could not have constructed a Lagrangian that both contained only first-order derivatives of the field (metric) quantities and was at the same time a scalar. He thus treated a more general case than Rosenfeld's, allowing for the Lagrangian to change by a total divergence under a general coordinate transformation.[4] Bergmann's derivation of the constraints was thus somewhat different from Rosenfeld's and went via the (generalized) contracted Bianchi identities. This actually lent a certain degree of coherence to Bergmann's overall approach, as the Bianchi identities also played an essential role in the generalized derivation of the EIH result.

Bergmann identified the identities (and the non-unique time development they implied) as the central challenge for the quantization of a generally covariant field theory. He was well aware that this problem could be remedied by the imposition of gauge, or rather, in this case, coordinate conditions, but general covariance was such a central concept to him that he instead aimed for constructing a Hamiltonian in which the arbitrariness of coordinate choice showed up explicitly, in the form of four arbitrary functions. He was thus aiming for a quantization procedure very alike that of Rosenfeld, yet without, it should be pointed out, being aware of Rosenfeld's work.

---

[3]Indeed, Bergmann's paper contains the cryptic remark that general relativity had "so far not successfully absorbed the existence of quantities possessing half-odd spin." It is an open historical problem what this statement refers to.

[4]A Lagrangian for General Relativity that contained only first-order derivatives of the metric had first been introduced by Einstein (1916), the so-called $\Gamma - \Gamma$-Lagrangian, given by Einstein originally only for a specific choice of coordinates. Rosenfeld had actually also considered Lagrangians that change by a total divergence, but only for "internal" symmetries that do not act on the space-time coordinates. It was thus not directly applicable to Lagrangians that changed by a total divergence under general coordinate transformations, but it did apply to Rosenfeld's tetrad Lagrangian, that changed by a total divergence only under local Lorentz transformations. Rosenfeld's tetrad Lagrangian was a scalar density under general coordinate transformations. His method can be straightforwardly extended to the case of coordinate transformations, as shown in Salisbury and Sundermeyer (2017).

The second paper he published in that same year, together with his colleague Johanna Brunings, represented a further step towards quantization (Chapter 33). The central point Bergmann wanted to address here was how to carry over the EIH method into quantum theory. The difficulty he had to address was the following: If there are singular world-lines in space-time, where the field equations are not satisfied, then one should not integrate over these singular regions in the action integral, whose minimization gives the field equations. However, without first solving the field equations, one cannot know the exact singular particle trajectories. Bergmann's solution was the "parameter formalism." The idea was that since the spacetime trajectories of the singularities could not be known in advance, a four-dimensional multiply connected parameter space (t, u^s) could be utilized in which parameter tubes could be excised without making a commitment as to the corresponding spacetime particle paths. The original spacetime coordinates x^ρ were taken to be functions of these parameters, thus becoming field variables on par with the gravitational field, dynamical equations of motion and all. The introduction of parameters was, as acknowledged by Bergmann, closely related to Weiss's approach discussed in Chapter 17: For Weiss, too, the field variables were no longer functions of the space-time coordinates, but rather of the parameters defining a point on the hypersurface, as well as of a further parameter (λ in Weiss's notation, t in Bergmann's), labelling a member of a sequence of hypersurfaces. And for Weiss the original space-time coordinates were also functions of the parameters. However, in Weiss's case this was merely to be understood as the definition of a specific foliation of space-time. As Bergmann pointed out, the emergence of the original coordinates as dynamical variables on the same footing as the field variables y^A, which was the essential point for Bergmann, was not discussed by Weiss, who consequently also did not investigate the momenta canonically conjugate to the original coordinates.

Bergmann's parameter formalism was now covariant not just under coordinate transformations but also under arbitrary reparameterizations. This covariance brought with it four new identities. Three of these (corresponding to the three spatial parameters u^s) could be straightforwardly deduced from the definition of the new canonical momenta conjugate to the now-dynamical spacetime coordinates. Repeating this procedure for the fourth (time) parameter t, did not directly deliver another constraint. Instead, Bergmann could show in this manner that the Lagrangian in the parameter formalism was homogeneous of degree one, in both the field velocities $\dot{y}_A$ and the coordinate velocities $\dot{x}^\rho$, where the dot refers to differentiation with respect to parameter time t in both cases. From this fact the existence of a further constraint equation $g = 0$ could be deduced.

This function g had the properties of the Hamiltonian density, that is, it generated the canonical field equations of motion through functional differentiation. Consequently, using Poisson brackets (in the parameter version, first introduced by Weiss in the 1930s, but now, in Bergmann and Brunings' approach, also including the space-time coordinates as dynamical variables) the spatial integral of g (the Hamiltonian) could be used to determine the time derivative of some functional of the canonical variables.[5] They then claimed to show, by calculating the Poisson bracket of the Hamiltonian with an arbitrary linear combination of the constraints, that all eight constraints had vanishing time derivatives, and

[5]Concerning these functionals, which were to represent physical quantities, Bergmann and Brunings made the crucial observation in this paper that not all functions of the original field variables and their first derivatives (velocities) possessed a counterpart in phase space. They identified the necessary and sufficient condition for converting a function in configuration-velocity-space into a phase-space function (such functions need to be, in modern parlance, constant along all of the null directions of the singular Legendre matrix) and dubbed quantities that satisfied this requirement "dynamical variables".

that consequently no further constraints had to be considered. As Bergmann himself would soon realize, this proof was fallacious.[6] They also outlined how to proceed to a quantum theory, in the same manner proposed by Rosenfeld, that is, by turning the Poisson brackets into commutators and the constraints into auxiliary conditions on the wave function.

But although several important properties of the constraint g were known, Bergmann and Brunings did not provide a method for explicitly constructing it. They *could* show that it was not unique: For example, if one added to g some linear combination of the remaining seven constraints, the resulting function would of course still be a constraint and could also still be taken as a Hamiltonian density, the new canonical field equations of motion being related to the ones of the original g by a suitable canonical transformation of the dynamical variables along with a transformation of the space-time parameters. This implied that some specific choice of Hamiltonian density corresponded to choosing a coordinate condition and thus losing general covariance. But Bergmann and Brunings neither showed how to construct a specific Hamiltonian, corresponding to a particular coordinate condition, nor a general Hamiltonian involving arbitrary functions. They thus had reached, in 1949, about the same point that Rosenfeld had reached in 1930. The next step was clearly the explicit construction of a Hamiltonian, only that now, in Bergmann's parameterized approach, this Hamiltonian would not just generate the field equations of motion and time evolution of physical variables, but would also itself be one of the constraints.[7]

In reaching this next step, Bergmann's group was beaten by Felix Pirani and Alfred Schild. Their work built on an approach to constrained Hamiltonian dynamics quite different from that of Rosenfeld or Bergmann, presented by Paul Dirac in August/September 1949 at the Summer Seminar of the Canadian Mathematical Congress (Chapter 34), a newly established graduate-level summer school, which in that year (its second installment) had a special focus on mathematical physics, other lecturers including Homi J. Bhabha (on quantum field theory) and Laurent Schwartz (on distribution theory; *Proceedings of the Second Canadian Mathematical Congress* 1951). Dirac came to constrained Hamiltonian dynamics from a direction quite different from that of Rosenfeld or Bergmann. He was not interested in the effect of symmetries when passing from a Lagrangian to a Hamiltonian formulation of a field theory, be they general local symmetries (Rosenfeld) or general covariance in particular (Bergmann). Dirac's focus was rather a subject that had been on his mind for quite some time: The relation between Hamiltonian dynamics and special relativity. Dirac viewed this as the central problem in constructing a (special) relativistic quantum (field) theory, a problem he still considered to be wide open, as he was not particularly inclined towards the renormalization methods being developed at the time.[8] He had already remarked in 1933:

> Quantum mechanics was built up on a foundation of analogy with the Hamiltonian theory of classical mechanics. This is because the classical notion of canonical coordinates and momenta was found to be one with a very simple

[6]This fact has already been pointed out by one of us (Salisbury 2006). It is hard to say precisely how Bergmann and Brunings came to this erroneous conclusion, since they do not give the details and there is no equation in the paper that is actually wrong. The erroneous statement is that "[t]he proof follows immediately if the functional $\mathscr{G}$ [the linear combination of primary constraints] is substituted in the expression for the Poisson bracket (3.23) […]" A simple counterexample is to be found in the toy model employed in Pitts (2014), which can easily be parameterized. We would like to thank James Brian Pitts for suggesting this check upon Bergmann and Brunings's claim.

[7]That the Hamiltonian of GR would be a constraint even in the unparameterized formulation, the so-called "problem of time," was only gradually realized in the course of the 1950s.

[8]See, e.g., Kragh (1990, 183).

> quantum analogue, as a result of which the whole of the classical Hamiltonian theory, which is just a structure built up on this notion, could be taken over in all its details into quantum mechanics. [...] [T]he Hamiltonian method is essentially non-relativistic in form, since it marks out a particular time variable as the canonical conjugate of the Hamiltonian function. (Dirac 1933b, 64.)

At the time, Dirac had believed that the way to proceed towards a relativistic quantum theory was consequently to find a new quantization procedure based on the classical Lagrangian, instead of the classical Hamiltonian, formalism. This was based on the realization that the action function appearing in the Lagrangian formalism, that is, the time integral of the Lagrangian itself, is a relativistic invariant. 16 years later, however, Dirac was considerably more optimistic about the reconcilability of Hamiltonian dynamics and relativity. The main reason for this appears to have been Dirac's adoption of Weiss's parameter formalism, which had been devised by Weiss precisely in order to obtain a covariant generalization of classical Hamiltonian dynamics, including canonical variables and Poisson brackets, as a starting point for the quantization of relativistic field theories.

In his Canadian lectures, Dirac went significantly beyond Weiss. Dirac emphasized that the parameter approach would turn the original space-time coordinates into dynamical variables in their own right. This came from a long tradition in Dirac's thinking: Already in the mid-1920s (Dirac 1926), and then again in 1933 (Dirac 1933a), Dirac had explored the possibility of turning time from an evolutionary parameter into a dynamical variable in relativistic particle mechanics, an approach which is the particle-mechanical analog of the field-theoretical parameter formalism. As opposed to Bergmann, however, Dirac did not place any physical expectations in the newly dynamical nature of the coordinates: To him, this was merely a mathematical observation, one that ensured the (Lorentz) covariance of the scheme. In his 1949 lectures, Dirac further recognized that the use of the parameter formalism would imply the existence of constraints, including the Hamiltonian constraint, another insight going back to his early particle-mechanical exploration of the subject. This helped explain how the dilemma in the above quote was solved: The parameter formalism allowed a separation between (a) the Hamiltonian function giving the total energy or energy density (which now appeared as one canonical momentum among others, conjugate to the dynamical time coordinate) and (b) the generator of the equations of motion and the time evolution of physical quantities (which was now the task of the constraints, or, as Dirac had called them in 1933, the Hamiltonian equations).

It was on this basis that Dirac built up his approach to constrained dynamics. The structure he obtained was general enough to also accommodate constraints that did not arise from the use of parameters (and also to deal with a non-parameterized theory, involving only other constraints). Dirac mentioned this fact, even though he did not refer to any examples (back in 1933, however, he had mentioned the analogy with the constraints arising in electrodynamics), but Pirani and Schild, who were attending the conference, quickly picked up on the applicability of Dirac's approach to General Relativity. Still, it should be emphasized that to Dirac the constraints arising from the parameters were the essential ones; additional constraints could simply be added on, as further identically fulfilled relations between canonical coordinates and momenta. Where these additional constraints might come from was of no concern to Dirac at the time and he did not discuss symmetry properties at all.[9] This of course made Dirac's constrained dynamics look

[9] Joshua Goldberg (2005) has thus labelled Dirac's approach as algebraic (concerned with the constraints themselves), as opposed to the group theoretical (concerned with the underlying symmetries) approach of Bergmann (and also Rosenfeld).

quite different from that of Bergmann, and one might spend a whole paper discussing the differences and similarities between the two. In the following, we will just briefly focus on the specific difference that was important for the work of Pirani and Schild, namely, that Dirac's formalism provided an explicit construction principle for the Hamiltonian.

In working out the properties of their constraint g, Bergmann and Brunings had, quite naturally, considered it to be a function of canonical coordinates and momenta only, allowing for derivatives of the canonical coordinates with respect to the spatial parameters u, but not allowing for velocities, that is, derivatives of the canonical coordinates with respect to the time parameter t. However, the Hamiltonian density H is quite easily written down as a function of canonical coordinates, momenta, *and* velocities as

$$H = p_n \dot{q}_n - L(q_n, \dot{q}_n) \tag{31.1}$$

using Dirac's notation, where the q_n are all the canonical coordinates (including the original space-time coordinates), the p_n are the conjugate momenta and the dot denotes differentiation with respect to the time parameter t. Now normally this is just an intermediate step in constructing the canonical Hamiltonian, the next step being the elimination of the velocities with the help of the defining equations for the momentum. This is, however, as we have already remarked in context of Rosenfeld's work in Chapter 17, a non-trivial task. It was now Dirac's innovation to consider Equation 31.1 not as an intermediate step—basically just a construction prescription—but rather as an actual expression for the Hamiltonian defined not on the usual phase space, but rather on a larger, $3N$-dimensional (with N the number of configuration variables) space, which also included the N velocities as independent variables. The usual Hamiltonian was then obtained by going to a $2N$-dimensional subspace, where all the defining equations for the momenta (including primary constraints) were fulfilled (this subspace of course corresponds to usual phase space if there are no constraints).

Dirac was now able to find another expression for the Hamiltonian density as a function of coordinates, momenta, and velocities:

$$H = v_m(p, q, \dot{q})\phi_m(p, q) \tag{31.2}$$

where the v_m are some suitable functions, the important thing being that out of each summand one of the constraints ϕ_m (the summation over m is over all constraints, those arising from parameterization and those intrinsic to the theory) can be factored out, the constraint being a function of coordinates and momenta alone.

Pirani and Schild, both in Toronto at the time, Schild working with Leopold Infeld and Pirani just having acquired his Master's degree, attended Dirac's Vancouver lectures in 1949. According to Pirani's recollections,[10] Schild immediately realized the applicability of Dirac's method to the quantization of general relativity.[11] Indeed, the paper of Bergmann and Brunings had appeared just a month before Dirac's lecture, and it is quite

[10] Interview by DR, 23 June 2011, https://www.aip.org/history-programs/niels-bohr-library/oral-histories/34463 (accessed 21 July 2017).

[11] Dirac had apparently not realized this and Schild brought it to his attention at the time. His wife, Winnie Schild, told one of the authors (DS) of her husband's pride in having achieved Dirac's enduring respect through this observation. More importantly, the fact that Dirac had not made the connection underlines that it was not as obvious as Pirani recalled it to be in the interview with DR, making Schild's previous exposure to Bergmann and Brunings's work all the more probable.

probable that Schild had seen it and then in Vancouver realized how Dirac's method might supply the Hamiltonian and the eighth constraint. He discussed the matter with Pirani, who promptly agreed to go with him to Pittsburgh (where Schild had just accepted a job at the Carnegie Institute of Technology) and become Schild's first PhD student, the subject of his thesis being the quantization of general relativity. Pirani's thesis (Pirani 1951) (which dealt, among other things, also with the inclusion of fermionic spinors) was not completed until 1951, but Pirani and Schild had already presented first results at the APS Meeting in New York (February 2–4, 1950, Schild and Pirani 1950) and submitted the paper reprinted here in Chapter 35 on February 14. By comparing the two forms of the Hamiltonian provided by Dirac, they were quite easily able to construct the first Hamiltonian for parameterized GR and thereby to also explicitly identify the eight primary constraints. Bergmann and his group were not far behind and published their Hamiltonian within the same year, constructed using a different procedure based on finding the "quasi-inverse" of the singular Legendre matrix (P. Bergmann et al. 1950).[12]

The success of Pirani and Schild and the Bergmann group in constructing a Hamiltonian for GR was initially seen as invalidating an approach based on the notion of a free graviton field, perturbed by its self-interaction and its interactions with other fields. Such an attempt at quantizing gravity had been worked out in the 1949 PhD thesis of Bryce DeWitt (né Seligman), a student of Schwinger's, using the newly developed covariant techniques of his supervisor (DeWitt 1949). DeWitt wrote up a paper based on his thesis[13] in 1950 in which he duly acknowledged "the rigorous Pirani-Schild-Dirac scheme," having been able to read a manuscript of the Pirani-Schild paper before publication. Still, DeWitt's paper, which he submitted to the Physical Review, was essentially rejected by the referee, H. P. Robertson. In his referee report,[14] he stated

> In view of the Pirani-Schild paper, which Seligman has seen—and remarks on p. 3, footnote, that this should eventually be carried out in terms of their more rigorous theory—it would seem to me better to suggest he carry out his work in terms of their theory.

DeWitt's thesis work was never published, and he indeed turned to working on the Pirani-Schild formulation soon after (DeWitt 1952;DeWitt and DeWitt-Morette 1952).

However, as it turned out, the gravitational Hamiltonians of 1950 were not the last word on quantum gravity, with only details left to be filled out. The most immediate foundational reform was the abandonment of the parameter formalism by both groups within the same year. For Bergmann the parameter formalism had carried the hope of a quantum version of EIH, in which the space-time coordinates, which were dynamical variables in the parameter formalism, could be interpreted as the coordinates of a quantum mechanical point particle. That this will not work is rather obvious from a modern perspective: The parameterization introduces four additional constraints which reduce the number of dynamical degrees of freedom by exactly the number of dynamical variables one had formally gained by introducing the parameters in first place. But recall that Bergmann and Brunings had originally convinced themselves that there were no secondary constraints in their formalism; they were simply ignoring four of the constraints actually present in the

[12] For more on the mathematical theory (and a brief history) of the generalized inverses of singular matrices, see Rao and Mitra (1971). In the nomenclature adopted there, the Bergmann group was working with a (not uniquely defined) reflexive g-inverse of the Legendre matrix.

[13] A manuscript is in the possession of his widow, Cecile Morette-DeWitt.

[14] 17 May 1950, H. P. Robertson Papers, Caltech

formalism. And indeed, Bergmann's abandonment of the parameter formalism appears to coincide with his realization that there were in fact additional secondary constraints.

In an interview, Joshua Goldberg placed the abandonment of the parameter formalism in the summer of 1950 or 1951,[15] with the published evidence clearly favoring the former (the first parameter-free paper by the Bergmann group was already submitted in May 1951, see Anderson and P. G. Bergmann (1951)). And in a short note to Schild, dated 16 November 1950,[16] Bergmann announced that his group had discovered additional constraints (i.e., had discovered the secondary constraints) and was working on a proof that there were as many additional constraints as there were primary ones. Such a proof was supplied a year later in the non-parameterized formalism (Anderson and P. G. Bergmann 1951), and really only makes sense therein, as there are more primary than secondary constraints in the parameter formalism. So there is strong circumstantial evidence that Bergmann's abandonment of the parameter formalism was closely linked with the discovery of the secondary constraints, even though Bergmann never made that point explicitly in writing. In fact, he never even retracted his "no secondaries" argument in his published work. In the first detailed, published presentation of the parameter-free formalism (Penfield 1951), the new approach was merely presented as equivalent to the parameter formalism (for singularity-free gravitational fields), the manifest covariance of the latter thus ensuring the covariance of the former. And only another two years later, after the question of which were to be the actual physical observables in a constrained theory moved to the center of Bergmann's attention, did he go on record stating that the parameterization did not, as he had hoped, introduce new observables (i.e., particle trajectories) into the theory (P. G. Bergmann and Schiller 1953).

The idea of introducing particle trajectories as new quantum operators via parameterization had never played a prominent role for Pirani and Schild. Parameterization had rather been a means (inherited from Dirac) for making the general covariance of the setup manifest, as it allowed for a definition of (parameter) time that did not presuppose any metric structure for space-time (which of course was to be a result of the field equations), but could be defined solely through the introduction of a family of three-dimensional hypersurfaces in the metric-free (Pirani used the term "amorphous" in his thesis) parameter space. When Schild received the above-mentioned note by Bergmann in November 1950, he had not yet considered abandoning parameters.[17] He and Pirani had (probably following Bergmann and Brunings) been convinced that there were no secondary constraints, without providing any argument for this statement. After receiving Bergmann's note, Pirani attempted to calculate the secondary constraints in the parameterized formalism. This proved to be a formidable task, too formidable indeed: Pirani, in his thesis, merely presented some preliminary calculations in this direction, stating (p. 42) that "[t]he long and complicated expressions which make their appearances […] led to a search for some way of simplifying the theory." Pirani and Schild thus followed the lead of the Bergmann

[15] Interview by DR and DS, 21 March 2011, https://www.aip.org/history-programs/niels-bohr-library/oral-histories/34461 (accessed 21 July 2017).

[16] Alfred Schild Papers, Dolph Briscoe Center for American History, The University of Texas at Austin.

[17] In the Schild Papers, there is a letter from Pirani dated 22 January 1951, in which the use of the parameter-free formalism is already discussed as a matter of course, seemingly contradicting the above statement. However, in this letter Pirani makes reference to a radio broadcast by Ralph Williamson, which was only broadcast in June 1951 (Williamson 1951), strongly implying a redating of this letter to 1952, making this an instance of the classic mistake of still dating with the old year in January.

group, dropping the parameters, merely viewing them as a proof of the general covariance of the Hamiltonian formulation (Pirani 1951).[18]

Later work on the canonical quantization of gravity thus looked quite different from that done in these early works of 1949/50. In particular, as pointed out by Pirani to Schild in a letter from 4 December 1952,[19] the problem of integrating over singular regions in the action integral, which Bergmann had avoided by introducing parameters, remained. Indeed the parameter-free formalism strongly implied entirely doing away with the idea of particles as singularities (and thus the whole quantum EIH approach), and instead adopting more modern, (quantum) field-theoretical descriptions of matter. Still, this early work provided a basis for all later attempts at canonical quantization, introducing most of the concepts and terminology regarding the appearance of constraints.

References

Anderson, James L. and Peter G. Bergmann (1951). Constraints in Covariant Field Theories. *Physical Review* 83:1018–1025.

Bergmann, Peter G., Robert Penfield, Ralph Schiller, and Henry Zatzkis (1950). The Hamiltonian of the General Theory of Relativity with Electromagnetic Field. *Physical Review* 80:81–88.

Bergmann, Peter G. and Ralph Schiller (1953). Classical and Quantum Field Theories in the Lagrangian Formalism. *Physical Review* 89:4–16.

DeWitt, Bryce (Dec. 1949). *I: The Theory of Gravitational Interactions. II: The Interaction of Gravitation with Light*. Published as Carl Bryce Seligman. PhD thesis. Harvard.

— (1952). Point Transformations in Quantum Mechanics. *Physical Review* 85:653–661.

DeWitt, Bryce and Cécile DeWitt-Morette (1952). The Quantum Theory of Interacting Gravitational and Spinor Fields. *Physical Review* 87:116–122.

Dirac, Paul A. M. (1926). Relativity Quantum Mechanics with an Application to Compton Scattering. *Proceedings of the Royal Society, Series A* 111(758):405–423.

— (1933a). Homogeneous Variables in Classical Dynamics. *Mathematical Proceedings of the Cambridge Philosophical Society* 29:389–400.

— (1933b). The Lagrangian in Quantum Mechanics. *Physikalische Zeitschrift der Sowjetunion* 3(1):64–72.

Einstein, Albert (1916). Die Grundlage der allgemeinen Relativitätstheorie. *Annalen der Physik* 49:769–822.

Goldberg, Joshua (2005). Syracuse: 1949–1952. In: *Einstein Studies Volume 11: The Universe of General Relativity*. Ed. by A. J. Kox and Jean Eisenstaedt. Basel: Birkhäuser, 357–371.

Halpern, Paul (2005). Peter Bergmann: The Education of a Physicist. *Physics in Perspective* 7:390–403.

Kragh, Helge (1990). *Dirac: A Scientific Biography*. Cambridge: Cambridge University Press.

Penfield, Robert (1951). Hamiltonians without Parameterization. *Physical Review* 84: 737–743.

[18] The resultant simplifications also allowed Pirani and Schild to identify a mistake in their first paper (see the January 1951/52 letter from Pirani to Schild mentioned in footnote 17): Equation 62 is wrong, no matter where one chooses to close the square bracket: If it is closed at the end of the expression, the sign of the last summand should be changed from plus to minus. If it is closed before the last summand, then that summand needs an additional factor of $l_\gamma l^{-2}$.

[19] Alfred Schild Papers

Pirani, Felix A. E. (1951). *On the Quantization of the Gravitational Field of General Relativity*. PhD thesis. Carnegie Institute of Technology.

Pitts, J. Brian (2014). Change in Hamiltonian General Relativity from the Lack of a Time-Like Killing Vector Field. *Studies in History and Philosophy of Modern Physics* 47: 68–89.

Proceedings of the Second Canadian Mathematical Congress (1951). Toronto: University of Toronto Press.

Rao, C. Radhakrishna and Sujit Kumar Mitra (1971). *Generalized Inverse of Matrices and its Applications*. New York: Wiley.

Salisbury, Donald (2006). Peter Bergmann and the invention of constrained Hamiltonian dynamics. *arXiv:physics/0608067v1*.

Salisbury, Donald and Kurt Sundermeyer (2017). Translation into English of Léon Rosenfeld's "Zur Quantelung der Wellenfelder" with commentary. *European Physical Journal H* 42(1):23–61.

Schild, Alfred and Felix A. E. Pirani (1950). On the Quantization of Einstein's Gravitational Field Equations. *Physical Review* 78:329.

Schweber, Silvan S. (1994). *QED and the Men Who Made It: Dyson, Feynman, Schwinger, and Tomonaga*. Princeton: Princeton University Press.

Williamson, Ralph E. (1951). Fred Hoyle's Universe. *The Journal of the Royal Astronomical Society of Canada* 45:185–189.

Chapter 32
Peter G. Bergmann (1949): Non-Linear Field Theories

Peter G. Bergmann (1949). Non-Linear Field Theories. *Physical Review*, 75: 680–685.

PHYSICAL REVIEW VOLUME 75, NUMBER 4 FEBRUARY 15, 1949

Non-Linear Field Theories

PETER G. BERGMANN
Department of Physics, Syracuse University, Syracuse, New York

(Received June 8, 1948)

This is the first paper in a program concerned with the quantization of field theories which are covariant with respect to general coordinate transformations, like the general theory of relativity. All these theories share the property that the existence and form of the equations of motion is a direct consequence of the covariant character of the equations. It is hoped that in the quantization of theories of this type some of the divergences which are ordinarily encountered in quantum field theories can be avoided. The present paper lays the classical foundation for this program: It examines the formal properties of covariant field equations, derives the form of the conservation laws, the form of the equations of motion, and the properties of the canonical momentum components which can be introduced.

I. INTRODUCTION

At the present time, two great theoretical structures in physics can lay claim to containing significant parts of the "truth" which to unearth must remain the principal aim of both the experimental and the theoretical physicist. One of these structures is modern quantum physics as applied to both mechanical and field theoretical problems; the other is the general theory of relativity, which in the author's opinion represents the least imperfect "classical" (i.e., non-quantized) field theory. It is well known that at this time each one of these structures suffers from serious and apparently inherent weaknesses; Quantum physics in most of its realizations still requires a two-stage development, an underlying classical theory (a mechanical system or a set of field equations together with equations of motion) followed by a process of quantization. The result of this quantization process, when applied to a combination of fields and particles, usually leads to characteristic divergences. The general theory of relativity, on the other hand, has so far not successfully absorbed the existence of quantities possessing half-odd spin, nor can it be quantized in a satisfactory manner; as a result, the theory has been completely useless in atomic and nuclear physics.

Neither theory has as yet made any significant contribution to the problem of the constitution of the elementary particles. Moreover, while each structure undeniably contains elements of truth, their combination has so far proved unsuccessful.

The purpose of the present program is to analyze each of the two theories for its essential and, presumably, relatively permanent contributions to our present knowledge and, thus, to construct what might be called skeletonized theories. An attempt will be made to see whether such a covariant theory is at all susceptible to quantization and whether the result will be an improved theory.

Specifically, it is believed that the theory of relativity contains two great permanent achievements: (a) it is the only theory of gravitation which explains reasonably the equality of inertial and gravitational mass (the so-called principle of equivalence); (b) it is the only classical field theory in which the equations of motion of particles in the field are contained in the field equations, instead of being logical juxtapositions. That is why it is possible, in the general theory of relativity, to treat the motion of field singularities (which are used to represent particles) without having to deal with infinite interaction terms of one kind or another. It is possible that this accomplishment will also lead to a more satisfactory quantized theory, although the author has little hope that such a theory would be sufficiently powerful to attack the problem of the constitution of elementary particles.

On the other hand, there is probably no particular reason why the theory of relativity must appear in the form of Riemannian geometry, i.e., using primarily the differential covariants of a symmetric tensor of rank 2. The first step in the program sketched above consists therefore, in the setting up of a generalized classical theory and the examination of its properties. This is the scope of the present paper.

The theory to be developed is essentially non-linear, and its Hamiltonian non-quadratic. It is impossible to envisage a set of differential equations which is covariant with respect to general coordinate transformations, yet linear. While this might appear to be merely an annoying complication, the non-linear character of the general theory of relativity is crucial for the possibility of interaction between different particles. It is possible to set up linear, Lorentz-covariant field equations for a field similar to gravitational potentials, which contain equations of motion; but these turn out to represent simply Newton's first law: Each singularity is tied inexorably to a straight-line uniform motion, without any interaction between pairs of particles. It is the non-linear terms in the field equations which provide for forces, gravitational and otherwise.[1-3]

II. THE FIELD VARIABLES AND THE LAGRANGIAN

Instead of introducing a metric tensor, as is done in the general theory of relativity, we shall leave the exact nature of the field variables unspecified, denoting them merely by the symbol y_A $(A=1, \cdots, N)$, where N is the number of algebraically independent components. It will be assumed that the field equations can be derived from a variational principle of the form

$$\delta I=0, \quad I=\int_V L(y_A, y_{A,\mu})d^4x. \qquad (2.1)$$

In other words, the field equations will determine values of the integral I which are stationary with respect to infinitesimal changes in the field variables; these changes are arbitrary, except that they must remain confined to the interior of the chosen four-dimensional volume V. The Lagrangian L is assumed to be an algebraic function of the field variables themselves and their first partial derivatives with respect to the coordinates. As is customary, $y_{A,\mu}$ is short for $\partial y_A/\partial x^\mu$.

The field equations which result from the variation of the field variables in the interior shall be designated by L^A,

$$L^A \equiv (\partial L/\partial y_A) - (\partial L/\partial y_{A,\rho})_{,\rho}=0. \qquad (2.2)$$

More precisely, the symbol L^A will be used for the left-hand side of these equations, no matter whether they are satisfied or not.

In this paper it will be assumed that the transformation law of the field variables is linear and homogeneous in the field variables themselves and that it depends algebraically on the first derivatives of the new coordinates with respect to the old ones. With respect to *infinitesimal* coordinate transformations, the field variables will transform according to a law having the form

$$\bar{\delta} y_A = F_{A\mu}{}^{B\nu}\xi^\mu{}_{,\nu} y_B - y_{A,\mu}\xi^\mu. \qquad (2.3)$$

The four functions ξ^μ represent the infinitesimal changes of the coordinate values of a fixed world point. The $F_{A\mu}{}^{B\nu}$ are numbers, independent of both the choice of coordinate system and the coordinate values themselves, but characteristic for the type of field variables representing the field. Finally, the $\bar{\delta} y_A$ are the changes produced in the field variables y_A *as functions of their arguments* because of the infinitesimal coordinate transformations. Because the transformed y_A are not compared with the original values at the same world point, but with the original values at that world point which possesses the same coordinate values prior to the transformation, Eq. (2.3) contains a "transport" term, the second term on the right-hand side. Incidentally, in this equation, as well as throughout the paper, the summation convention is being applied both to the Greek indices which are associated with the coordinates and run from 1 to 4 and to the capital indices which run from 1 to N.

The transformation law (2.3) is, of course, not the most general law which may be encountered in geometrical objects, but it does include all types of tensors and tensor densities and also spinors. Being an infinitesimal transformation law, it is subject to the requirement that the commutator should again be an operator of the same type. This condition is represented by the

[1] Einstein, Infeld, and Hoffman, Annals of Mathematics **39**, 65 (1938).

[2] A. Einstein and L. Infeld, Annals of Mathematics **41**, 455 (1940).

[3] L. Infeld and P. R. Wallace, Phys. Rev. **57**, 797 (1940).

identity

$$F_{A\mu}{}^{C\nu}F_{C\rho}{}^{B\sigma}-F_{A\rho}{}^{C\sigma}F_{C\mu}{}^{B\nu}$$
$$\equiv\delta_\rho{}^\nu F_{A\mu}{}^{B\sigma}-\delta_\mu{}^\sigma F_{A\rho}{}^{B\nu}. \quad (2.4)$$

The field equations (2.2) must be covariant, i.e., if they are satisfied in one coordinate system, they must also be satisfied automatically in any other coordinate system. They would certainly be covariant if the integral I were an invariant, for in that case its extremization would also be an invariant operation. However, this condition is too strong, as it cannot even be satisfied in the general theory of relativity. There it is possible to introduce either a Lagrangian of which the integral is an invariant but which contains explicitly second derivatives of the field variables, or one which contains only first derivatives but which has no invariant integral. If the Lagrangian, in the face of an infinitesimal coordinate transformation, adds a divergence, then that condition is sufficient (though possibly not necessary) to assure covariant field equations. We shall, therefore, require that

$$\bar{\delta}L=Q^\mu{}_{,\mu}, \quad (2.5)$$

where the four expressions Q^μ are some functions of the ξ^μ and their derivatives (including those of higher order). In other words, it will be assumed that the infinitesimal change in the integral I can be expressed by means of a surface integral

$$\bar{\delta}I=\oint_S Q^\mu n_\mu dS, \quad (2.6)$$

without being affected by the values of the ξ^μ in the interior. The four quantities n_μ represent the components of a "unit normal vector" which is introduced so that Gauss' theorem can be formulated. This is possible even without the introduction of a metric.

By considering the transformation of a variation of the Lagrangian, $\bar{\delta}(\delta L)$, it can easily be shown that the field equations, because of Eq. (2.5) or its equivalent (2.6), satisfy the transformation law

$$\bar{\delta}L^B=-F_{A\mu}{}^{B\nu}\xi^\mu{}_{,\nu}L^A-(L^B\xi^\mu)_{,\mu}, \quad (2.7)$$

where the constants $F_{A\mu}{}^{B\nu}$ are identical with those introduced in Eq. (2.3).

III. IDENTITIES

Consider again the transformation law (2.5), which specifies that in the event of an infinitesimal coordinate transformation the change in the Lagrangian is a divergence. That change can also be represented in the form of a *variation* induced by the infinitesimal transformation:

$$\bar{\delta}L=L^A\bar{\delta}y_A+((\partial L/\partial y_{A,\rho})\bar{\delta}y_A)_{,\rho}. \quad (3.1)$$

Substitution of Eqs. (2.3) and (2.5) results in

$$(Q^\rho-(\partial L/\partial y_{A,\rho})\bar{\delta}y_A)_{,\rho}$$
$$=L^A(F_{A\mu}{}^{B\nu}\xi^\mu{}_{,\nu}y_B-y_{A,\mu}\xi^\mu). \quad (3.2)$$

The right-hand side will be a divergence only if the left-hand sides of the field equations satisfy the four identities

$$(F_{A\mu}{}^{B\nu}y_BL^A)_{,\nu}+y_{A,\mu}L^A\equiv 0. \quad (3.3)$$

In the general theory of relativity these identities are known as the contracted Bianchi identities. They hold no matter whether the field equations are satisfied or not.

The expressions L^A contain the field variables themselves and also their first and second derivatives. The second derivatives occur only linearly, and their coefficients can be represented in the form

$$\begin{aligned} L^A&=L^{AB\rho\sigma}y_{B,\rho\sigma}+\cdots,\\ L^{AB\rho\sigma}&=-\tfrac{1}{2}[(\partial^2L/\partial y_{A,\rho}\partial y_{B,\sigma})\\ &\qquad+(\partial^2L/\partial y_{A,\sigma}\partial y_{B,\rho})]. \end{aligned} \quad (3.4)$$

When the L^A are substituted into the identities (3.3), the terms containing second derivatives will, in turn, lead to terms containing third derivatives, and those must cancel each other, irrespective of other terms,

$$F_{A\mu}{}^{B\rho}y_BL^{AC\sigma\tau}y_{C,\rho\sigma\tau}\equiv 0. \quad (3.5)$$

It follows that the coefficients $L^{AB\sigma\tau}$ must satisfy the identities

$$(F_{A\mu}{}^{B\rho}L^{AC\sigma\tau}+F_{A\mu}{}^{B\sigma}L^{AC\tau\rho}$$
$$+F_{A\mu}{}^{B\tau}L^{AC\rho\sigma})y_B\equiv 0. \quad (3.6)$$

For what follows, $4N$ of these identities are of special interest, those in which ρ, σ, and τ all equal 4,

$$F_{A\mu}{}^{B4}\Lambda^{AC}y_B \equiv 0,$$
$$\Lambda^{AC} = L^{AC44} = -\partial^2 L/\partial y_{A,4}\partial y_{C,4}. \tag{3.7}$$

Before concluding this section, it will be well also to formulate the conservation laws which are satisfied in a covariant theory. It is well known that in the presence of a Lagrangian, and provided the field equations are satisfied, there exist 16 quantities $t_\iota{}^\kappa$, functions of the field variables and their first derivatives only, which satisfy four divergence relationships:

$$t_\iota{}^\rho{}_{,\rho} = 0,$$
$$t_\iota{}^\kappa = \delta_\iota{}^\kappa L - y_{A,\iota}(\delta L/\delta y_{A,\kappa}). \tag{3.8}$$

If now the presence of matter in the field is represented by continuous and differentiable right-hand sides of the field equations,

$$L^A = P^A, \tag{3.9}$$

then the right-hand side of Eq. (3.8) no longer vanishes, but one obtains instead

$$t_\iota{}^\rho{}_{,\rho} = y_{A,\iota}P^A. \tag{3.10}$$

Because of the existence of the differential identities (3.3), this right-hand side can also be given the form of a divergence. Naturally, the field equations (3.9) can be satisfied only if the right-hand sides satisfy the same relationships which are satisfied identically by the left-hand sides. Therefore, Eq. (3.10) may be written in the form

$$T_{\iota,\rho}{}^\rho = 0,$$
$$T_\iota{}^\kappa = \delta_\iota{}^\kappa L - y_{A,\iota}(\partial L/\partial y_{A,\kappa}) + F_{A\iota}{}^{B\kappa}P^A y_B. \tag{3.11}$$

Only in this "strong" form are the conservation laws useful in the consideration of situations in which matter is represented by continuous expressions P^A, or else by discrete singularities of the field variables.

IV. EQUATIONS OF MOTION

The equations of motion are obtained by the method initiated by Einstein and collaborators.[1,2] Because of the identities derived in Section III, any singularities present in the field are subject to certain restrictions which represent both the conservation of mass (or its equivalent) and the equations of motion. Just as in those papers, it is impossible to formulate these conditions precisely without taking recourse to an *approximation method.* This is because the motions of the singularities cannot be completely determined unless such effects as spontaneous polarization and spontaneous emission of radiation by the singularities are specifically excluded. This exclusion is accomplished by assuming that all motions are "slow" in the sense that differentiation of a field variable with respect to x^4 (the time) reduces the order of magnitude at every stage of the approximation method.

It shall be assumed that solutions of the field equations are to be obtained in the form of a power series expansion with respect to some parameter ϵ, which might, for instance, represent the order of magnitude of the material velocities involved, (v/c):

$$y_A = \overset{0}{y}_A + \epsilon\overset{1}{y}_A + \epsilon^2\overset{2}{y}_A + \cdots. \tag{4.1}$$

Moreover, the zeroth approximation shall be the "trivial" solution, a rigorous solution of the field equations in which all field variables are constants. (In the theory of relativity, this trivial solution is represented by the flat Minkowski metric.) The first approximation must then satisfy the following linear, homogeneous equations:

$$\overset{0}{L}{}^{ABrs}\overset{1}{y}_{B,rs} = 0. \tag{4.2}$$

In this approximation, there appear no time derivatives, because of the assumption that these are of a higher order of magnitude. The $\overset{0}{L}{}^{ABrs}$ are the coefficients (3.4), and the indices r and s are coordinate indices running from 1 to 3, in accordance with the usual notation in the literature. Because of the condition of "slow motion," this approximation will not contain radiation, but rather solutions corresponding to material particles, with as yet undetermined motions.

Generally, these solutions will not be defined throughout space. There will be singularities, probably of the $(1/r)$ type; more precisely, at each instant of time (x^4) there will be certain three-dimensional domains in which the field equations have no bounded solutions. However, only such solutions will be considered in which each one of these singular regions can be surrounded by a closed surface S (in three-dimensional space, S is two dimensional) on which the field equations are satisfied. It is this condition of separateness of the singular regions which makes possible the formulation of additional requirements on the solutions outside.

If we proceed to the next approximation, the equations will have the form

$$\overset{0}{L}{}^{ABrs}\overset{2}{y}_{B,rs} = -\overset{2}{L}{}^{A}(\overset{0}{y}+\epsilon \overset{1}{y}). \tag{4.3}$$

The right-hand sides will be the L^A of the second order, formed from the first-order solutions (the $\overset{1}{L}{}^A$ vanish, of course). They will contain terms which are linear in the first and second time derivatives of the first approximation, and quadratic in the purely spatial derivatives of the first approximation.

Because of Eqs. (3.6), the left-hand sides of the second-order equations (4.3) satisfy identically the four relationships

$$F_{A\mu}{}^{Ct}\overset{0}{y}_C(\overset{0}{L}{}^{ABrs}\overset{2}{y}_{B,rs})_{,t} \equiv 0, \tag{4.4}$$

irrespective of the choice of the second-order field variables. In particular, it is possible to satisfy the field equations outside the singular regions and to continue the second-order variables throughout the interior of the singular regions with arbitrary continuous and three times differentiable functions. Then the conditions (4.4) will be satisfied throughout (three-dimensional) space and will permit the application of Gauss' theorem to a three-dimensional domain which includes a singular region but which is bounded by a closed surface on which the field equations are satisfied. On that surface, on which Eqs. (4.3) are to be satisfied, we have then

$$\begin{aligned} 0 &\equiv \oint_S F_{A\mu}{}^{Ct}\overset{0}{y}_C\overset{0}{L}{}^{ABrs}\overset{2}{y}_{B,rs}n_t dS \\ &= -\oint_S F_{A\mu}{}^{Ct}\overset{0}{y}_C\overset{2}{L}{}^{A}(\overset{0}{y}+\epsilon\overset{1}{y})n_t dS, \end{aligned} \tag{4.5}$$

which represent for that singularity inside S the three equations of motion and the law of the conservation of mass (μ takes all four values $1\cdots 4$). It can be shown easily that these conditions for the first approximation are empty unless S encloses a singular region. The number of conditions which can be obtained equals, therefore, four times the number of separate singular regions.

V. UNIQUENESS OF SOLUTIONS, MOMENTA

In this section it will be shown first that if the field variables and their first derivatives are given on a three-dimensional hypersurface in space-time (e.g., throughout three-space at a specified time), the continuation of the solution beyond that initial hypersurface is not unique. As a corollary, it will then be shown that if canonical momenta are introduced in that formalism, not all the time derivatives of the field variables can be expressed in terms of these momenta. Both of these results can be predicted qualitatively from the covariance of the field equations. Suppose that one solution were known which satisfies the initial conditions on the hypersurface. Then one could always carry out a coordinate transformation such that the new coordinates coincide on the surface with the old ones, up to second derivatives, but not elsewhere. Then the solutions would be transformed into *formally* different solutions satisfying the same initial conditions. *Physically*, one would be inclined to call such equivalent solutions "the same solution;" in that sense, the solutions are presumably uniquely determined by the initial conditions.

It was pointed out previously that the field equations are linear with respect to the second derivatives of the field variables. If we choose as the hypersurface one with x^4 constant, then the y_A as well as the $y_{A,4}$ are given on that hypersurface. The continuation would be unique if the field equations could be solved with respect to the terms containing $y_{A,44}$. This, however, is not possible. Consider the coefficients of these second time derivatives. They are the quantities L^{AB44}. If the matrix

$$\|\Lambda^{AB}\| = \|L^{AB44}\| \tag{5.1}$$

were regular, then this solution could be accomplished. Actually, there are four different sets of N quantities which are zero eigenvectors of that matrix, as shown in Eq. (3.7). Conversely, there exist four linear combinations of field equations which are free of second time derivatives and which, therefore, represent restrictions on the choice of initial conditions on the hypersurface. They are

$$y_B F_{A\mu}{}^{B4} L^A = 0. \tag{5.2}$$

It follows that four linear combinations of the second time derivatives remain completely arbitrary on each space-like hypersurface. Naturally, the singling out of some particular coordinate as the "time" is completely arbitrary, and the term "time" was used only to conform to physical intuition.

The ambiguous character of the continuation of the solution with given initial conditions leads to pecularities of the canonical momenta which are analogous to those encountered in quantum electrodynamics. As is customary, the derivatives of L with respect to $y_{A,4}$ are designated as the momenta,

$$\partial L/\partial y_{A,4} = \pi^A. \tag{5.3}$$

These N equations cannot be solved with respect to the $y_{A,4}$, but, on the contrary, there exist four relationships between the π^A and y_A, $y_{A,s}$. First, Eqs. (5.3) can be solved with respect to $y_{A,4}$ only if the determinant of the partial derivatives

$$(\partial\pi^A(y_B, y_{B,s}, y_{B,4})/\partial y_{C,4}) = (\partial^2 L/\partial y_{A,4}\partial y_{C,4}) = -\Lambda^{AC} \tag{5.4}$$

does not vanish. But Λ^{AC} is a singular matrix, and its determinant is zero. It follows that an attempt to solve with respect to the $y_{A,4}$ will result in the establishment of four relationships which do not contain any (first-order) time derivatives. In view of the $4N$ identities

$$y_B F_{A\mu}{}^{B4}(\partial\pi^A/\partial y_{C,4}) \equiv 0, \tag{5.5}$$

these four relationships can be obtained by straightforward integration. They are

$$y_B F_{A\mu}{}^{B4}\pi^A - K_\mu(y_C, y_{C,s}) \equiv 0, \tag{5.6}$$

where the K_μ are functions introduced by the integration, but actually determined in any theory.

VII. CONCLUSION

The relationships set up in the last Section will give rise to the usual difficulties in quantization, since the N momenta are not algebraically independent of each other. Four of the Hamiltonian equations will turn out to be empty, as a result. When the field variables y_A and π^A are reinterpreted as operators, it will not be possible to interpret Eqs. (5.6) as linear relationships satisfied by the operators π^A; such an assumption would be incompatible with the commutation relations. Rather, they will have to be interpreted as initial conditions which are imposed on the state vector and which are preserved automatically in the course of time. At each instant, the infinitesimal contact transformation leading from the state at t to the state at $(t+dt)$ will contain four arbitrary functions of the spatial coordinates. These arbitrary functions are usually eliminated by the setting up of so-called coordinate conditions, analogous to the gauge condition of electrodynamics; but it is also possible to retain this arbitrariness in the formalism and set up a quantized theory in which the Hamiltonian is determined only up to four arbitrary functions. It is proposed to examine these problems in a future paper.

Chapter 33
Peter G. Bergmann and Johanna H. M. Brunings (1949): Non-Linear Field Theories II: Canonical Equations and Quantization

Peter G. Bergmann and Johanna H. M. Brunings (1949). Non-Linear Field Theories II: Canonical Equations and Quantization. *Reviews of Modern Physics*, 21: 480–487.

REVIEWS OF MODERN PHYSICS VOLUME 21, NUMBER 3 JULY, 1949

Non-Linear Field Theories II. Canonical Equations and Quantization*

PETER G. BERGMANN AND JOHANNA H. M. BRUNINGS
Department of Physics, Syracuse University, Syracuse, New York

In this paper, a covariant field theory of the general type of the theory of relativity is brought into the canonical form and then quantized. Particles are assumed to be represented as singularities of the field. Primarily, we had to overcome two difficulties. First, the variational integral should be extended only over that space-time domain which is free of singularities. Since the location of the singular world lines cannot be known until after the integration of the field equations has been completed, we have introduced a second set of coordinates—called "parameters" in this paper—which will serve as variables of integration and in terms of which the motions of the particles can be arbitrarily prescribed. The second difficulty arises in that the expressions for the canonical momentum densities cannot be solved with respect to the partial time derivatives of the field variables; this circumstance precludes the construction of the Hamiltonian by the usual methods. Nevertheless, we have shown that a Hamiltonian exists, though it is not uniquely determined by the Lagrangian; our Hamiltonian contains an arbitrary linear combination of the eight algebraic relationships that exist between the canonical variables at each world point. The canonical field equations have the usual form. They are covariant if the choice of Hamiltonian is left open. The eight algebraic constraints on the canonical variables at each point are all integrals of the field equations. So are the Poisson brackets between the canonical variables (at the same time). When this system of equations and constraints is quantized, the property of general covariance can be used to carry out a proof of the covariance of the whole theory, including the commutation relations, that requires none of the computational effort usually required in theories that are merely Lorentz invariant. Once the system of equations has been completed, it turns out that the covariance goes much farther than was required originally. Because of the introduction of the parameters, the ordinary coordinates of space-time turn formally, at least, into dynamical variables, and the usual canonical transformations, with respect to which the theory is covariant, transform the coordinates, the original field variables, and the canonical conjugates of both together. The canonical conjugates of the coordinates are the expressions ordinarily interpreted as energy and momentum densities. The physical significance of these canonical transformations, which cause the world points to lose their identities, is not yet understood.

1. INTRODUCTION AND REVIEW

IN a previous paper,[1] attention was called to the special physical features of completely covariant field theories. If the field equations are the Euler-Lagrange equations of a four-dimensional variational principle, it is possible to obtain equations of motion for particles by considering the particles as singular time-like curves in space-time and by requiring that outside these singularities the field equations are satisfied everywhere. This dependence of the equations of motion on the field equations was first established for the general theory of relativity by A. Einstein and his co-workers.[2–4]

In I, it was assumed that the field equations can be derived from a variational principle which contains only the field variables and their first partial derivatives, but no higher derivatives, and which is so constructed that the resulting Euler-Lagrange equations are covariant. It was shown that the field equations satisfy identities similar to the Bianchi identities of the general theory of relativity, that they cannot be solved with respect to the highest second derivatives with respect to some one coordinate, and that the continuation of solutions which are given on an initial space-like hypersurface in the direction of time is not uniquely determined by the equations. It was further shown that in this generalized theory (in which the nature of the field remains for the time being unspecified) the equations of motion of field singularities are determined by the field equations outside in the same manner as in the general theory of relativity.

Because the motion of field singularities is determined by the field outside the singularities (if each singularity is enclosed in a small three-dimensional spherical surface) the field equations need to be satisfied only on and outside that spherical envelope. The classical (i.e., unquantized) theory can be carried through without considering the self-energy of the particles or any other divergent quantities. To this extent, this type of theory is the most nearly self-consistent classical field theory yet devised. It still falls short of the "perfect" field theory in that it fails to represent particles as non-singular solutions of the field equations. The purpose of our present program is to attempt the quantization of such a field theory and to see to what extent the usual divergences of quantum field theories can be avoided. The present paper is a step in this program. We shall first of all indicate how a variational principle can be set up which requires integration only over the non-singular domain of space-time. We shall then develop the canonical form of the classical theory, and we shall show finally that if the classical field variables are replaced by operators, then the usual commutation relations are covariant.

For the convenience of the reader, we shall collect

* This work was supported in part by the Office of Naval Research under Contract N6-onr-248.

[1] P. G. Bergmann, Phys. Rev. 75, 680 (1949). Referred to in this paper as I. Reference to formulas in that paper is made as (I-4.2), etc.

[2] A. Einstein, L. Infeld, and B. Hoffmann, Ann. Math. 39, 65 (1938).

[3] A. Einstein and L. Infeld, Ann. Math. 41, 455 (1940).

[4] L. Infeld and P. R. Wallace, Phys. Rev. 57, 797 (1940).

here those results and formulas of I which are needed in this paper. We denote the field variables by $y_A (A=1, \cdots, N)$, and their derivatives with respect to one of the four coordinates x^ρ by $y_{A,\rho}$. We shall assume that the Lagrangian L depends only on y_A and on $y_{A,\rho}$. The field equations then take the form

$$0 = \partial^A L - (\partial^{A\rho} L)_{,\rho} \equiv L^A \qquad \text{(I-2.2)}$$

where the symbol ∂^A denotes differentiation of a function of the field variables with respect to y_A and the symbol $\partial^{A\rho}$ differentiation with respect to $y_{A,\rho}$. An infinitesimal coordinate transformation is defined by means of the four functions ξ^ρ, which represent the (infinitesimal) changes in the values of the coordinates of a fixed world point with the original coordinates x^ρ. It is assumed that the transformation law of the field variables with respect to infinitesimal coordinate transformations possesses the form

$$\bar{\delta} y_A = F_{A\mu}{}^{B\nu} \xi^\mu{}_{,\nu} y_B - y_{A,\mu} \xi^\mu. \qquad \text{(I-2.3)}$$

The coefficients $F_{A\mu}{}^{B\nu}$ are a set of constants characteristic for the field variables. The last term in (I-2.3) is a "transport term," which must be inserted if the value of y_A at a point with the original coordinates $\mathbf{x}$ is to be compared with the transformed y_A at the point which *after the transformation* possesses the coordinates $\mathbf{x}$ (and, therefore, originally had the coordinates $\mathbf{x} - \boldsymbol{\xi}$).

The transformation matrix $F_{A\mu}{}^{B\nu}$ must satisfy certain commutation relations if the infinitesimal transformation law is to generate a finite transformation law.

$$F_{A\mu}{}^{C\nu} F_{C\rho}{}^{B\sigma} - F_{A\rho}{}^{C\sigma} F_{C\mu}{}^{B\nu} \equiv \delta_\rho{}^\nu F_{A\mu}{}^{B\sigma} - \delta_\mu{}^\sigma F_{A\rho}{}^{B\nu}. \qquad \text{(I-2.4)}$$

The Lagrangian from which the field equations are to be derived is assumed to satisfy a transformation law of the form

$$\bar{\delta} L = Q^\rho{}_{,\rho} \qquad \text{(I-2.5)}$$

which assures that the left-hand sides of the field equations transform covariantly according to the formula

$$\bar{\delta} L^B = -F_{A\mu}{}^{B\nu} \xi^\mu{}_{,\nu} L^A - (L^B \xi^\mu)_{,\mu}. \qquad \text{(I-2.7)}$$

It can easily be shown that the Q^ρ of (I-2.5) are the following four expressions:

$$Q^\rho \equiv F_{A\mu}{}^{B\rho} y_B L^A \xi^\mu + \partial^{A\rho} L \bar{\delta} y_A. \qquad (1.1)$$

The most important property of the field equations for what follows is that they satisfy four differential identities: When we substitute for L^A the full expressions in terms of the field variables and their first and second derivatives, the four equations

$$F_{A\mu}{}^{B\rho} (L^A y_B)_{,\rho} + L^A y_{A,\mu} \equiv 0 \qquad \text{(I-3.3)}$$

are identically satisfied. These equations contain the third derivatives of the y_A linearly. Their coefficients must vanish identically by themselves. These conditions turn out to be

$$\begin{aligned}(F_{A\mu}{}^{B\rho} L^{AC\sigma\tau} + F_{A\mu}{}^{B\sigma} L^{AC\tau\rho} + F_{A\mu}{}^{B\tau} L^{AC\rho\sigma}) y_B \equiv 0,& \\ L^{AC\rho\sigma} = -\tfrac{1}{2}(\partial^{A\rho}\partial^{C\sigma} L + \partial^{A\sigma}\partial^{C\rho} L).& \qquad \text{(I-3.6)}\end{aligned}$$

Finally, if the field equations are satisfied, then the divergence of certain expressions vanishes, which are usually identified with the energy momentum densities and energy-momentum fluxes:

$$\begin{aligned} t_\iota{}^\rho{}_{,\rho} &= 0 \\ t_\iota{}^\kappa &= \delta_\iota{}^\kappa L - y_{A,\iota} \partial^{A\kappa} L. \qquad \text{(I-3.8)}\end{aligned}$$

In these formulas and all that follow, the summation convention for dummy indices is used for all types of indices without distinction.

2. THE PARAMETER FORMALISM

In his formulation of the equations of motion in the general theory of relativity, Einstein worked exclusively with the field equations and with the identities that exist between them without taking recourse to the Lagrangian. In attempting to construct a Hamiltonian as a preliminary to quantization, we shall need to refer to the original variational principle. The volume integral in space-time $\int L d\mathbf{x}$ that is to be made stationary with respect to variations of the field variables in the interior of the domain of integration can be extended over any four-dimensional domain desired, and the field equations will then be satisfied throughout the interior of that domain. Naturally, the domain of integration should not include regions where the equations cannot be satisfied, and it should in particular exclude the world lines along which particles move.

In this connection, there arises a peculiar difficulty. Because the motions of the particles are determined by the field equations outside, we cannot predict the location of the singular world lines until after we have accomplished the integration of the field equations. On the other hand, it will, of course, be necessary to choose the domain of integration for the Lagrangian even before the field equations are formulated, let alone solved. This difficulty can be resolved if we introduce, in addition to the coordinates x^ρ, a second set of coordinates that can serve as variables of integration. Call the second set $u^s (s=1, \cdots, 3)$, t. To avoid any ambiguities of language, this second set will be referred to as "parameters," and the one parameter t will occasionally be called "time." The transformation law leading from the coordinates to the parameters shall not be specified initially, except that it shall be assumed that the Jacobian of the transformation,

$$J \equiv \left| \frac{\partial(x^\rho)}{\partial(u^s, t)} \right|, \qquad (2.1)$$

shall not vanish. In terms of parameters, the basic variational principle goes over into

$$\begin{aligned} \delta S &= 0 \\ S &= \int_t \int_{\mathbf{u}} J L d\mathbf{u} dt. \qquad (2.2)\end{aligned}$$

The coordinates x^ρ are considered as field variables, along with the y_A. Naturally, the additional field equations will not lead to a determination of the x^ρ as functions of the parameters if the field equations are to be satisfied in a singly connected four-dimensional domain. Straightforward computation shows that the Euler-Lagrange equations belonging to the new field variables x^ρ reduce to (I-3.8) and are, thus, contained in the original field equations. In the presence of line singularities, however, the coordinates of the singularities are, in fact, determined by certain integrals over the (space-like, two-dimensional) closed surfaces that enclose the singularities and which are derivable from the divergence relations (I-3.8) or from the stronger relations (I-3.11), which hold even inside the singular regions and which go over into (I-3.8) on the outside.

Nothing prevents us from prescribing the motions of the singularities in terms of the parameters *initially* and from extending the integral (2.2) over the multiply connected domain of the four parameters in which we propose to satisfy the field equations. If the integral is made stationary with respect to variations of the field variables y_A, x^ρ in the interior of that domain, both the field equations and the equations of motion will be satisfied.

The formal device just described permits us to formulate our variational principle so that the integrand of S is everywhere finite. Naturally, the field equations must be covariant both with respect to coordinate transformations and with respect to parameter transformations. With respect to the latter, the field variables are scalars and their first derivatives covariant vectors. In what follows, all $(N+4)$ field variables will be accepted as such.

Since partial derivatives will occasionally have to be taken, both with respect to x^ρ and with respect to u^s and t, we shall uniformly denote differentiation with respect to t by the dot symbol and differentiation with respect to u^s by the symbol $_{|s}$. We have, thus, for example,

$$y_{A,\rho}=\dot{y}_A t_{,\rho}+y_{A|s}u^s{}_{,\rho} \tag{2.3}$$

and

$$\begin{aligned}\dot{y}_A&=y_{A,\rho}\dot{x}^\rho\\ y_{A|s}&=y_{A,\rho}x^\rho{}_{|s}.\end{aligned} \tag{2.4}$$

3. CONONICAL VARIABLES AND THE HAMILTONIAN

Into the formalism just described canonical momentum densities can be introduced in the usual manner by means of the defining equations,

$$\pi^A\equiv\partial^{A\cdot}(LJ)=Jt_{,\rho}\partial^{A\rho}L \tag{3.1}$$

and

$$\begin{aligned}\lambda_\rho\equiv\partial(LJ)/\partial\dot{x}^\rho&=Jt_{,\sigma}(L\delta_\rho{}^\sigma-y^A{}_{,\rho}\partial^{A\sigma}L)\\ &\equiv Jt_{,\sigma}t_\rho{}^\sigma.\end{aligned} \tag{3.2}$$

$t_\rho{}^\sigma$ has been defined by means of Eq. (I-3.8). These $(N+4)$ canonical momentum densities satisfy a number of algebraic relations. With the help of Eq. (I-3.6), four of these relations can be obtained. By multiplying (I-3.6) by $J^2\cdot t_{,\rho}t_{,\sigma}t_{,\tau}$ one obtains, after a few obvious calculations,

$$\partial^{A\cdot}(F_{B\mu}{}^{C\tau}y_C\pi^B Jt_{,\tau})=0. \tag{3.3}$$

The expression within the parentheses of (3.3) could depend on $\dot{x}^\rho$ either directly or through the $y_{A,\rho}$. The latter possibility is excluded because

$$\left.\begin{aligned}(\partial y^A{}_{,\sigma}/\partial\dot{x}^\rho)\partial^{A\sigma}&=-y_{A,\rho}t_{,\sigma}\partial^{A\sigma}\\ &=-y_{A,\rho}\partial^{A\cdot}\end{aligned}\right\}. \tag{3.4}$$

But the same expression in (3.3) cannot depends on $\dot{x}^\rho$ directly, either, because

$$\frac{\partial(Jt,\sigma)}{\partial\dot{x}^\rho}=0. \tag{3.5}$$

It follows that integration of the four Eqs. (3.3) leads to the four equations

$$F_{A\mu}{}^{B\tau}y_B\pi^A Jt_{,\tau}=K_\mu(y_C,y_{C|s},x^\rho{}_{|s}), \tag{3.6}$$

where the functions on the right hand side are determined in any actual theory and, in any case, do not contain either $\dot{y}_A$ or $\dot{x}^\rho$ as arguments.

Additional algebraic relationships can be obtained directly from the defining Eqs. (3.1), (3.2). By multiplying the obvious identity

$$\lambda_\rho-Jt_{,\rho}L+y_{A,\rho}\pi^A=0 \tag{3.7}$$

by $x^\rho{}_{|s}$, we find

$$x^\rho{}_{|s}\lambda_\rho+y_{A|s}\pi^A=0, \tag{3.8}$$

three algebraic relationships between the canonical variables which are free of "time" derivatives. If we multiply Eq. (3.7) by $\dot{x}^\rho$, we get

$$JL\equiv\dot{y}_A\pi^A+\dot{x}^\rho\lambda_\rho, \tag{3.9}$$

showing formally that the Lagrangian JL is homogeneous of the first degree in the arguments $\dot{y}_A$, $\dot{x}^\rho$. If the Lagrangian JL is of the first degree, then the derivatives π^A and λ_ρ are homogeneous of the zeroth degree. It is well known that n zeroth-degree functions of n arguments cannot be algebraically independent of each other. It is possible to obtain a differential relationship for this algebraic condition. Let the condition be

$$g(\pi^A,\lambda_\rho,y_A,y_{A|s},x^\rho{}_{|s})\equiv 0 \tag{3.10}$$

to be satisfied identically if the momenta are replaced by their defining equations. Differentiation leads to

$$\begin{aligned}0=&\frac{\partial g}{\partial\pi^A}\delta\pi^A+\frac{\partial g}{\partial\lambda_\rho}\delta\lambda_\rho+\frac{\delta g}{\delta y_A}\delta y_A+\frac{\delta g}{\delta x^\rho}\delta x^\rho\\ &+\left(\frac{\partial g}{\partial y_{A|s}}\delta y_A+\frac{\partial g}{\partial x^\rho{}_{|s}}\delta x^\rho\right)_{|s}=\left[\frac{\partial g}{\partial\pi^B}\partial^{A\cdot}\partial^{B\cdot}(LJ)\right.\\ &\left.+\frac{\partial g}{\partial\lambda_\sigma}\partial^{A\cdot}\partial_{\sigma\cdot}(LJ)\right]\delta\dot{y}_A+\left[\frac{\partial g}{\partial\pi^B}\partial_{\rho\cdot}\partial^{B\cdot}(LJ)\right.\end{aligned}$$

$$
\begin{aligned}
&+\frac{\partial g}{\partial \lambda_\sigma}\partial_\rho.\partial_\sigma.(LJ)\Big]\delta\dot{x}\rho+\Big[\frac{\partial g}{\partial \pi^B}\frac{\partial \pi^B}{\partial y_A}+\frac{\partial g}{\partial \lambda_\sigma}\frac{\partial \lambda_\sigma}{\partial y_A} \\
&-\left(\frac{\partial g}{\partial \pi^B}\frac{\partial \pi^B}{\partial y_{A|s}}+\frac{\partial g}{\partial \lambda_\sigma}\frac{\partial \lambda_\sigma}{\partial y_{A|s}}\right)_{|s}+\frac{\delta g}{\delta y_A}\Big]\delta y_A \\
&-\left(\frac{\partial g}{\partial \pi^B}\frac{\partial \pi^B}{\partial x^\rho{}_{|s}}+\frac{\partial g}{\partial \lambda_\sigma}\frac{\partial \lambda_\sigma}{\partial x^\rho{}_{|s}}+\frac{\partial g}{\partial x^\rho{}_{|s}}\right)_{|s}\delta x^\rho \\
&+\Big[\left(\frac{\partial g}{\partial y_{A|s}}+\frac{\partial g}{\partial \pi^B}\frac{\partial \pi^B}{\partial y_{A|s}}+\frac{\partial g}{\partial \lambda_\sigma}\frac{\partial \lambda_\sigma}{\partial y_{A|s}}\right)\delta y_A \\
&\qquad+\left(\frac{\partial g}{\partial .x^\rho{}_{|s}}+\frac{\partial g}{\partial \pi^B}\frac{\partial \pi^B}{\partial x^\rho{}_{|s}}+\frac{\partial g}{\partial \lambda_\sigma}\frac{\partial \lambda_\sigma}{\partial x^\rho{}_{|s}}\right)\delta x^\rho\Big]_{|s}. \qquad (3.11)
\end{aligned}
$$

If we integrate this differential form over a space-like domain (with t constant) and if we consider the δy_A, $\delta\dot{y}_A$, etc., as arbitrary variations of the independent field variables, we get the conditions

$$
\begin{aligned}
&\frac{\partial g}{\partial \pi^B}\partial^{A\cdot}\partial^{B\cdot}(JL)+\frac{\partial g}{\partial \lambda_\sigma}\partial^{A\cdot}\partial_\sigma.(JL)=0, \\
&\frac{\partial g}{\partial \pi^B}\partial_\rho.\partial^{B\cdot}(JL)+\frac{\partial g}{\partial \lambda_\sigma}\partial_\rho.\partial_\sigma.(JL)=0, \\
&\frac{\partial g}{\partial y_A}+\frac{\partial g}{\partial \pi^B}\frac{\partial \pi^B}{\partial y_A}+\frac{\partial g}{\partial \lambda_\sigma}\frac{\partial \lambda_\sigma}{\partial y_A}=0, \\
&\frac{\partial g}{\partial y_{A|s}}+\frac{\partial g}{\partial \pi^B}\frac{\partial \pi^B}{\partial y_{A|s}}+\frac{\partial g}{\partial \lambda_\sigma}\frac{\partial \lambda_\sigma}{\partial y_{A|s}}=0, \\
&\frac{\partial g}{\partial x^\rho{}_{|s}}+\frac{\partial g}{\partial \pi^B}\frac{\partial \pi^B}{\partial x^\rho{}_{|s}}+\frac{\partial g}{\partial \lambda_\sigma}\frac{\partial \lambda_\sigma}{\partial x^\rho{}_{|s}}=0. \qquad (3.12)
\end{aligned}
$$

Equations (3.12) are, of course, satisfied by any one of the seven algebraic "constraints" on the canonical momenta (3.6), (3.8) as well as by the eighth relationship yet to be obtained. The first two Eqs. (3.12) can be understood in the sense that the $(N+4)$ quantities $(\partial g/\partial\pi^A)$, $(\partial g/\partial\lambda_\rho)$ form a "null vector" of the (singular) matrix

$$
\Lambda=\begin{pmatrix}\partial^{A\cdot}\partial^{B\cdot}(JL), & \partial^{A\cdot}\partial_\sigma.(JL)\\ \partial_\rho.\partial^{B\cdot}(JL), & \partial_\rho.\partial_\sigma.(JL)\end{pmatrix}. \qquad (3.13)
$$

This matrix has eight linearly independent null vectors. Seven of these null vectors will lead back to the relationships already obtained. An eighth null vector possesses the components $\dot{y}_A$, $\dot{x}^\rho$. The expressions

$$
\left.\begin{aligned}
\dot{y}_B\partial^{B\cdot}\partial^{A\cdot}(JL)+\dot{x}^\sigma\partial_\sigma.\partial^{A\cdot}(JL)=0\\
\dot{y}_B\partial^{B\cdot}\partial_\rho.(JL)+\dot{x}^\sigma\partial_\sigma.\partial_\rho.(JL)=0
\end{aligned}\right\} \qquad (3.14)
$$

vanish, because of the homogeneity property of the Lagrangian JL. We set, therefore,

$$
\begin{aligned}
&\frac{\partial g}{\partial \pi^A}=\dot{y}_A, \quad \frac{\partial g}{\partial \lambda_\rho}=\dot{x}^\rho, \\
&\frac{\delta g}{\delta y_A}=-\left(\dot{y}_B\frac{\partial \pi^B}{\partial y_A}+\dot{x}^\sigma\frac{\partial \lambda_\sigma}{\partial y_A}\right) \\
&\qquad\qquad+\left(\dot{y}_B\frac{\partial \pi^B}{\partial y_{A|s}}+\dot{x}^\sigma\frac{\partial \lambda_\sigma}{\partial y_{A|s}}\right)_{|s}\equiv-\frac{\delta(JL)}{\delta y_A}, \\
&\frac{\delta g}{\delta x^\rho}=-\frac{\delta(JL)}{\delta x^\rho}. \qquad (3.15)
\end{aligned}
$$

If the field Eqs. (I-2.2) are satisfied, the partial (or rather variational) derivatives of the Lagrangian with respect to y_A and x^ρ satisfy the equations

$$
\left.\begin{aligned}
\frac{\delta(JL)}{\delta y_A}=\dot{\pi}^A\\
\frac{\delta(JL)}{\delta x^\rho}=\dot{\lambda}_\rho
\end{aligned}\right\} \qquad (3.16)
$$

As a result, we have

$$
\dot{\pi}^A=-\frac{\delta g}{\delta y_A}, \quad \dot{\lambda}_\rho=-\frac{\delta g}{\delta x^\rho}, \quad \dot{y}_A=+\frac{\partial g}{\partial \pi^A}, \quad \dot{x}^\rho=+\frac{\partial g}{\partial \lambda_\rho}, \qquad (3.17)
$$

the canonical equations, with the function g as the Hamiltonian density.

The function g is not completely determined by the Eqs. (3.15). Suppose we have obtained an algebraic constraint g which is a result of the homogeneity of the right hand sides of Eqs. (3.1), (3.2) with respect to the $\dot{y}_A$, $\dot{x}^\rho$. Then first of all, this constraint can be multiplied by any non-vanishing function ν_0 of the x^ρ without destroying the validity of the Eqs. (3.15). All we do by such a change in g is to choose a new parameter system, (u^s,t^*), in which t^* is a function of the parameters (u^s,t) of the original system. Furthermore, if we add to some chosen g a linear combination of the remaining seven constraints, the coefficients being arbitrary functions of the $\mathbf{u},t$, we shall obtain a new set of canonical equations which differs from the original set in that the coordinates and parameters occurring in the new set are related to the coordinates and parameters of the original set of canonical equations by specific coordinate and parameter transformations.

We shall call a particular choice of the function g the Hamiltonian density and designate that choice by H. Once H has been chosen, the "equations of motion" (in the sense in which this term is used in quantum mechanical literature) are no longer covariant. Covariance can be maintained only as long as we are willing to reserve the choice of H.

In addition to the Hamiltonian density H, we shall introduce the Hamiltonian itself.

$$\mathcal{H} = \int_{\mathbf{u}} H du^1 du^2 du^3. \tag{3.18}$$

$\mathcal{H}$ is a functional and depends on the values of all the canonical variables throughout a whole space-like three-dimensional surface $t=$constant. "Functional derivatives" of a functional with respect to its arguments are to be understood in accordance with the defining equation

$$\delta\mathfrak{F} = \int_{\mathbf{u}} \left[\frac{\mathfrak{d}\mathfrak{F}}{\mathfrak{d}y_A(\mathbf{u})}\delta y_A(\mathbf{u}) + \frac{\mathfrak{d}\mathfrak{F}}{\mathfrak{d}\pi^A(\mathbf{u})}\delta\pi^A(\mathbf{u}) + \frac{\mathfrak{d}\mathfrak{F}}{\mathfrak{d}x^\rho(\mathbf{u})}\delta x^\rho(\mathbf{u}) + \frac{\mathfrak{d}\mathfrak{F}}{\mathfrak{d}\lambda_\rho(\mathbf{u})}\delta\lambda_\rho(\mathbf{u}) \right] \mathbf{du}, \tag{3.19}$$

where $\mathfrak{F}$ is any functional which depends on the canonical variables and their space derivatives throughout a three-dimensional surface $t=$constant. For an integral like the Hamiltonian, the functional derivatives reduce to the variational derivatives of the integrand,

$$\frac{\mathfrak{d}\mathcal{H}}{\mathfrak{d}y_A(\mathbf{u})} = \left.\frac{\delta H}{\delta y_A}\right|_{\mathbf{u}}, \text{ etc.} \tag{3.20}$$

With the help of the Hamiltonian $\mathcal{H}$ and its functional derivatives, the canonical equations can be rewritten in the form

$$\dot y_A = \frac{\mathfrak{d}\mathcal{H}}{\mathfrak{d}\pi^A(\mathbf{u})}, \quad \dot\pi^A = -\frac{\mathfrak{d}\mathcal{H}}{\mathfrak{d}y_A(\mathbf{u})}$$
$$\dot x^\rho = \frac{\mathfrak{d}\mathcal{H}}{\mathfrak{d}\lambda_\rho(\mathbf{u})}, \quad \dot\lambda_\rho = -\frac{\mathfrak{d}\mathcal{H}}{\mathfrak{d}x^\rho(\mathbf{u})}. \tag{3.21}$$

If we introduce a functional $\mathfrak{F}$ which depends on the canonical variables on one surface $t=$constant, then we can summarize Eqs. (3.21) in the form

$$\dot{\mathfrak{F}} = (\mathfrak{F}, \mathcal{H}) \tag{3.22}$$

where the right hand side, the "Poisson bracket" between $\mathfrak{F}$ and $\mathcal{H}$, stands for

$$(\mathfrak{F},\mathcal{H}) \equiv \int_{\mathbf{u}} \left\{ \frac{\mathfrak{d}\mathfrak{F}}{\mathfrak{d}y_A(\mathbf{u})}\frac{\mathfrak{d}\mathcal{H}}{\mathfrak{d}\pi^A(\mathbf{u})} - \frac{\mathfrak{d}\mathfrak{F}}{\mathfrak{d}\pi^A(\mathbf{u})}\frac{\mathfrak{d}\mathcal{H}}{\mathfrak{d}y_A(\mathbf{u})} + \frac{\mathfrak{d}\mathfrak{F}}{\mathfrak{d}x^\rho(\mathbf{u})}\frac{\mathfrak{d}\mathcal{H}}{\mathfrak{d}\lambda_\rho(\mathbf{u})} - \frac{\mathfrak{d}\mathfrak{F}}{\mathfrak{d}\lambda_\rho(\mathbf{u})}\frac{\mathfrak{d}\mathcal{H}}{\mathfrak{d}x^\rho(\mathbf{u})} \right\} \mathbf{du}. \tag{3.23}$$

It should be noted that not every functional of the original field variables y_A and their first derivatives $y_{A,\rho}$ can be converted into a functional of the canonical variables. If Ψ is a functional of y_A, $\dot y_A$, x^ρ, $\dot x^\rho$, then it can be represented as a functional of the canonical variables only if

$$\frac{\mathfrak{d}\Psi}{\mathfrak{d}\dot y_A} = \frac{\mathfrak{d}\Psi}{\mathfrak{d}\pi^B}\frac{\partial\pi^B}{\partial\dot y_A} + \frac{\mathfrak{d}\Psi}{\mathfrak{d}\lambda_\sigma}\frac{\partial\lambda_\sigma}{\partial\dot y_A} = \frac{\mathfrak{d}\Psi}{\mathfrak{d}\pi^B}\partial^{B\cdot}\partial^{A\cdot}(JL) + \frac{\mathfrak{d}\Psi}{\mathfrak{d}\lambda_\sigma}\partial_{\sigma\cdot}\partial^{A\cdot}(JL)$$
$$\frac{\mathfrak{d}\Psi}{\mathfrak{d}\dot x^\rho} = \frac{\mathfrak{d}\Psi}{\mathfrak{d}\pi^B}\frac{\partial\pi^B}{\partial\dot x^\rho} + \frac{\mathfrak{d}\Psi}{\mathfrak{d}\lambda_\sigma}\frac{\partial\lambda_\sigma}{\partial\dot x^\rho} = \frac{\mathfrak{d}\Psi}{\mathfrak{d}\pi^B}\partial^{B\cdot}\partial_{\rho\cdot}(JL) + \frac{\mathfrak{d}\Psi}{\mathfrak{d}\lambda_\sigma}\partial_{\sigma\cdot}\partial_{\rho\cdot}(JL). \tag{3.24}$$

In other words, the "dot products" of the "vector" $(\mathfrak{d}\Psi/\mathfrak{d}\dot y_A, \mathfrak{d}\Psi/\mathfrak{d}\dot x^\rho)$ by any one of the eight null vectors of the matrix Λ, Eq. (3.13), must vanish. With $\mathfrak{G}$ some linear functional of the eight algebraic constraints on the canonical variables, Eq. (3.24) can be rewritten in the form

$$\int_{\mathbf{u}} \left[\frac{\mathfrak{d}\Psi}{\mathfrak{d}\dot y_A(\mathbf{u})}\frac{\mathfrak{d}\mathfrak{G}}{\mathfrak{d}\pi^A(\mathbf{u})} + \frac{\mathfrak{d}\Psi}{\mathfrak{d}\dot x^\rho(\mathbf{u})}\frac{\mathfrak{d}\mathfrak{G}}{\mathfrak{d}\lambda_\rho(\mathbf{u})} \right] \mathbf{du} = 0. \tag{3.25}$$

Functionals which satisfy the condition (3.25) for all constraints and which can, therefore, be considered to depend only on the canonical variables shall be called "dynamical variables."

The functionals $\mathfrak{G}$ themselves are not only dynamical variables, they are "integrals of the motion": If the constraints are satisfied on one surface $t=$constant, then they remain satisfied for other values of t, provided the Eqs. (3.21) or (3.22) are satisfied. The proof follows immediately if the functional $\mathfrak{G}$ is substituted in the expression for the Poisson bracket (3.23) and if the expressions for $\mathfrak{d}\mathfrak{G}/\mathfrak{d}y_A$ and for $\mathfrak{d}\mathfrak{G}/\mathfrak{d}x^\rho$ are substituted from Eqs. (3.12).

4. POISSON BRACKETS AND CANONICAL TRANSFORMATIONS

It is appropriate in the canonical formalism to consider functionals and their time dependence in preference to functions. Most of the calculations become simplified if they are carried out on functionals, if we are only willing to assume that the two dimensional surface integrals which appear in integrations by parts give zero contributions. Ordinary functions may be considered as functionals whose functional derivatives with respect to canonical variables at points other than the one specified space point vanish.

As a preparation for the quantization of the theory, we shall in this section define the canonical transformations so that with respect to them the Poisson brackets of two dynamical variables are invariant, and we shall show that the canonical equations generate a canonical transformation.

Since we shall treat from now on the y_A and x^ρ uniformly, and likewise the $(N+4)$ canonically conjugate momentum densities, all formulas will be simplified if we combine the y_A and x^ρ into y_a, $a=1,\cdots,N+4$, and likewise the π^A and λ_ρ into π^a. Small Latin indices from the first half of the alphabet will now always run from 1 to $N+4$. With this notation, the Poisson bracket between two arbitrary dynamical variables $\mathcal{E}$ and $\mathcal{F}$ is defined as

$$(\mathcal{E},\mathcal{F})=\int_{\mathbf{u}}\left[\frac{\mathfrak{d}\mathcal{E}}{\mathfrak{d}y_a(\mathbf{u})}\frac{\mathfrak{d}\mathcal{F}}{\mathfrak{d}\pi^a(\mathbf{u})}-\frac{\mathfrak{d}\mathcal{E}}{\mathfrak{d}\pi^a(\mathbf{u})}\frac{\mathfrak{d}\mathcal{F}}{\mathfrak{d}y_a(\mathbf{u})}\right]\mathbf{du}. \quad (4.1)$$

Our dynamical system is completely defined by the requirements that for all values of t

$$\dot{\mathcal{F}}=(\mathcal{F},\mathcal{H})+\frac{\partial\mathcal{F}}{\partial t} \quad (4.2)$$

and that for $t=t_0$ the constraints

$$\mathcal{G}=0. \quad (4.3)$$

Now we can introduce a new system of canonical variables $y_a'(\mathbf{u}')$, $\pi^{a'}(\mathbf{u}')$ in such a manner that the system of Eqs. (4.2) goes over into a new system possessing the same form. In the usual manner, we shall introduce a "generating functional."

$$\mathcal{C}=\mathcal{C}(y_a(\mathbf{u}),\pi^{a'}(\mathbf{u}'),t) \quad (4.4)$$

and set

$$\left.\begin{aligned}\pi^a(\mathbf{u})&=\frac{\mathfrak{d}\mathcal{C}}{\mathfrak{d}y_a(\mathbf{u})}\\ y_a'(\mathbf{u}')&=\frac{\mathfrak{d}\mathcal{C}}{\mathfrak{d}\pi^{a'}(\mathbf{u}')}\\ \mathcal{H}'&=\mathcal{H}+\frac{\partial\mathcal{C}}{\partial t}\end{aligned}\right\}. \quad (4.5)$$

The identical transformation is generated by the functional

$$\left.\begin{aligned}\mathcal{C}_{(E)}&=\int_{\mathbf{u}}\int_{\mathbf{u}'}y_a(\mathbf{u})\pi^{b'}(\mathbf{u}')\delta_b{}^a\delta(\mathbf{u}'-\mathbf{u})\mathbf{du}'\mathbf{du}\\ &=\int_{\mathbf{u}}y_a(\mathbf{u})\pi^{a'}(\mathbf{u})\mathbf{du}\end{aligned}\right\}. \quad (4.6)$$

The infinitesimal transformation belonging to the transformation (4.4), (4.5) is, therefore,

$$\left.\begin{aligned}&\mathcal{C}=\mathcal{C}_{(E)}+c(y_a(\mathbf{u}),\pi^{a'}(\mathbf{u}),t),\\ &\delta y_a=\frac{\mathfrak{d}c}{\mathfrak{d}\pi^a},\quad \delta\pi^a=-\frac{\mathfrak{d}c}{\mathfrak{d}y_a},\quad \delta\mathcal{H}=\frac{\partial c}{\partial t}\end{aligned}\right\}. \quad (4.7)$$

Now it is easy to prove that with respect to the infinitesimal transformation the Poisson brackets do not change. We have

$$\left.\begin{aligned}\delta(\mathcal{E},\mathcal{F})=\int_{\mathbf{u}}\Bigg[&\delta\left(\frac{\mathfrak{d}\mathcal{E}}{\mathfrak{d}y_a}\right)\frac{\mathfrak{d}\mathcal{F}}{\mathfrak{d}\pi^a}+\frac{\mathfrak{d}\mathcal{E}}{\mathfrak{d}y_a}\delta\left(\frac{\mathfrak{d}\mathcal{F}}{\mathfrak{d}\pi^a}\right)\\ &-\delta\left(\frac{\mathfrak{d}\mathcal{E}}{\mathfrak{d}\pi^a}\right)\frac{\mathfrak{d}\mathcal{F}}{\mathfrak{d}y_a}-\frac{\mathfrak{d}\mathcal{E}}{\mathfrak{d}\pi^a}\delta\left(\frac{\mathfrak{d}\mathcal{F}}{\mathfrak{d}y_a}\right)\Bigg]\mathbf{du}\end{aligned}\right\} \quad (4.8)$$

where

$$\begin{aligned}\delta\left[\frac{\mathfrak{d}\mathcal{E}}{\mathfrak{d}y_a(\mathbf{u})}\right]&=-\int_{\mathbf{u}'}\left[\frac{\mathfrak{d}\mathcal{E}}{\mathfrak{d}y_b(\mathbf{u}')}\frac{\mathfrak{d}\delta y_b(\mathbf{u}')}{\mathfrak{d}y_a(\mathbf{u})}+\frac{\mathfrak{d}\mathcal{E}}{\mathfrak{d}\pi^b(\mathbf{u}')}\frac{\mathfrak{d}\delta\pi^b(\mathbf{u}')}{\mathfrak{d}y_a(\mathbf{u})}\right]\mathbf{du}'\\ &=+\int_{\mathbf{u}'}\left[\frac{\mathfrak{d}\mathcal{E}}{\mathfrak{d}\pi^b(\mathbf{u}')}\frac{\mathfrak{d}^2c}{\mathfrak{d}y_a(\mathbf{u})\mathfrak{d}y_b(\mathbf{u}')}-\frac{\mathfrak{d}\mathcal{E}}{\mathfrak{d}y_b(\mathbf{u}')}\frac{\mathfrak{d}^2c}{\mathfrak{d}y_a(\mathbf{u})\mathfrak{d}\pi^b(\mathbf{u}')}\right]\mathbf{du}'\ \text{etc.}\end{aligned} \quad (4.9)$$

If the expressions (4.9) are substituted into Eq. (4.8), the right-hand side vanishes identically. Since the result obtained for the infinitesimal transformation can be extended immediately to the finite transformation, it follows that the Poisson brackets are invariant with respect to the canonical transformations (4.4), (4.5).

Returning to the "equations of motion," (4.2), we shall now show that the right-hand side is invariant with respect to an infinitesimal canonical transformation. Because of the invariance of the bracket operation as such, the transformation of Eq. (4.2) yields

$$\left.\begin{aligned}\delta\dot{\mathcal{F}}&=(\mathcal{F},\delta\mathcal{H})+\delta\left(\frac{\partial\mathcal{F}}{\partial t}\right)\\ \delta\mathcal{H}&\equiv\frac{\partial c}{\partial t}\end{aligned}\right\}. \quad (4.10)$$

The last term is necessary because of the assumed explicit time dependence of the transformation. It equals

$$\left.\begin{aligned}\delta\left(\frac{\partial\mathcal{F}}{\partial t}\right)&=-\int_{\mathbf{u}}\left[\frac{\mathfrak{d}\mathcal{F}}{\mathfrak{d}y_a(\mathbf{u})}\frac{\partial\delta y_a(\mathbf{u})}{\partial t}+\frac{\mathfrak{d}\mathcal{F}}{\mathfrak{d}\pi^a(\mathbf{u})}\frac{\partial\delta\pi^a(\mathbf{u})}{\partial t}\right]\mathbf{du}\\ &=\int_{\mathbf{u}}\left[\frac{\mathfrak{d}\mathcal{F}}{\mathfrak{d}\pi^a}\frac{\mathfrak{d}(\partial c/\partial t)}{\mathfrak{d}y_a}-\frac{\mathfrak{d}\mathcal{F}}{\mathfrak{d}y_a}\frac{\mathfrak{d}(\partial c/\partial t)}{\mathfrak{d}\pi^a}\right]\mathbf{du}\\ &=-\left(\mathcal{F},\frac{\partial c}{\partial t}\right).\end{aligned}\right\} \quad (4.11)$$

As a result, the right-hand side of (4.10) vanishes.

If the "time" derivatives of all dynamical variables are invariant with respect to canonical transformations, it follows that the constraints $\mathcal{G}$ remain integrals of the motion. Setting them equal to zero initially assures their being zero permanently. Herewith, the proof of the invariance of the system of canonical equations (4.2), (4.3) with respect to canonical transformations (4.4), (4.5) is completed.

It is very easy to show that the change of the canonical variables with time is nothing but a special canonical transformation. Comparison of Eq. (4.7) with (3.21) shows that the generating functional is the Hamiltonian itself. Because the equations which connect the values of the canonical coordinates at two different times t_1 and t_2 with each other are equivalent to a canonical transformation, it follows that if $\mathcal{E}$ and $\mathfrak{F}$ are two functionals whose values are independent of t (though they will in general depend explicitly on the canonical variables and on t), then their Poisson bracket $(\mathcal{E},\mathfrak{F})$ is again constant in time.

5. QUANTIZATION

In this section, we shall demonstrate the covariance of the quantized theory which is obtained from the "classical" (i.e., unquantized) theory in the usual manner.

As is customary, we shall consider the canonical variables and also all the dynamical variables (in accordance with the defining Eq. (3.24)) as Hermitian operators acting on a "wave function" Ψ. For any two Hermitian operators $\mathcal{E}$ and $\mathfrak{F}$ which are the analogues of dynamical variables in the classical theory we shall require that the commutator,

$$[\mathcal{E},\mathfrak{F}] \equiv -\frac{i}{h}(\mathcal{E}\mathfrak{F}-\mathfrak{F}\mathcal{E}) \tag{5.1}$$

be the operator that is the analogue to the Poisson bracket $(\mathcal{E},\mathfrak{F})$. This correspondence requirement can always be satisfied if the operators $\mathcal{E}$ and $\mathfrak{F}$ are built up from the canonical operators in the same manner as the analogous dynamical variables from the canonical variables. The sequence of canonical operators in products has to be arranged so that the resulting operator is Hermitian.

The dependence of the operators on t is given by the new equation of motion

$$\mathfrak{F} = [\mathfrak{F},\mathcal{H}] + \frac{\partial\mathfrak{F}}{\partial t} \tag{5.2}$$

and the commutation relations between the canonical variables at the time t are given by the requirement (5.1),

$$[y_a(\mathbf{u}),y_b(\mathbf{u}')]=0 \quad [\pi^a(\mathbf{u}),\pi^b(\mathbf{u}')]=0 \quad [y_a(\mathbf{u}),\pi^b(\mathbf{u}')]=\delta_a{}^b\delta(\mathbf{u}-\mathbf{u}'). \tag{5.3}$$

Integration of the equations of motion gives for any operator which does not depend on t explicitly the expression

$$\mathfrak{F}\Big|_t = \exp\left(\frac{i}{h}\int_{t_0}^{t}\mathcal{H}(\tau)d\tau\right)\mathfrak{F}\Big|_{t_0}\exp\left(-\frac{i}{h}\int_{t_0}^{t}\mathcal{H}(\tau)d\tau\right). \tag{5.4}$$

All the operators $\mathcal{G}$ commute with the Hamiltonian. We shall require that at some initial time t_0 every $\mathcal{G}$ satisfies the condition

$$\mathcal{G}\Psi=0. \tag{5.5}$$

That condition will then be satisfied permanently.

The equations of motion (5.2), the commutation relations (5.1) or (5.3), and the constraints (5.5) are together a complete (Heisenberg) representation of a quantized theory. We have now to show that this theory is covariant with respect to coordinate and with respect to parameter transformations. That covariance is, of course, contingent on the continued free choice of the Hamiltonian as an arbitrary linear combination of constraint conditions $\mathcal{G}$, provided only that the coefficient of the eighth condition, the one that expresses the homogeneity of the Lagrangian, is nowhere zero in the $(\mathbf{u},t)$ continuum. Transition to a new coordinate system or to a new parameter system will carry the equations of motion over into themselves, but in general the form of the Hamiltonian will change. The problem is whether the commutation relations and the constraints will be reproduced by a transformation.

To answer this question requires no computational work, because of the special structure of our theory. It turns out that the constraint conditions as well as the commutation relations are covariant simply because they are integrals of the motion. Consider a solution of the equations of motion which also satisfies the constraints and the commutation relations. Naturally, this solution is available in a particular coordinate system and in a particular parameter system. Now it is possible to envisage a particular point of the $(\mathbf{u},t)$ continuum and to inquire how the constraints as well as the commutation relations will be affected by a coordinate or by a parameter transformation. Such a transformation can be achieved solely by an integration of the equations of motion. We merely need to carry our solution in one direction of t for a finite distance and then to integrate it backward, but with a different choice of Hamiltonian. This switch in the Hamiltonian will result in our returning to the original space-time point with a different set of coordinates and parameters. Specifically, that addition of a functional that is linear in the constraints (3.6) will transform the coordinates, while the constraints (3.8) and the eighth constraint are related to $\mathbf{u}$-transformations and t-transformations, respectively. Since integration of the equations of motion with arbitrary $\mathcal{H}$ preserves both the form of the constraints and the form of the commutation relations, covariance of all these relationships is assured.

6. DISCUSSION

With this paper, we have succeeded in putting a covariant theory into such a form that its equations can be integrated even in the presence of (world line) singularities. Moreover, we have shown that a Hamiltonian exists even though the differential equations do not determine uniquely the solutions of the standard Cauchy problem. This lack of determinancy of the solutions is reflected in a partial freedom of choice for the Hamiltonian: It is permissible to add to a given Hamiltonian an arbitrary linear combination of the eight algebraic constraints that hold between the canonical variables at each world point, with the only restriction that in the resulting new Hamiltonian the coefficient of the eighth constraint must not vanish.** Adoption of a particular linear combination of the constraints as the Hamiltonian amounts to the choice of a particular set of coordinate-parameter conditions. No matter what the choice of Hamiltonian, all other constraints $\mathcal{G}$ as well as the commutators of the canonical variables commute with the Hamiltonian and are, therefore, constant in time.

Our parameter formalism has formal similarity with one developed by P. Weiss.[5] It differs in that Weiss uses his analogue primarily to vary the domain of integration of the variational principle. His surfaces $t=$constant are necessarily closed, while ours are not only open, but, in the presence of particles, multiply connected. Also, Weiss has apparently taken the emergence of the x^{ρ} as dynamical variables not as seriously as we do.

Our formalism is covariant not only with respect to coordinate and parameter transformations, but also with respect to the much more general group of the canonical transformations (4.4), (4.5). Once the coordinates x^{ρ} and the energy-momentum densities λ_{ρ} are included in the dynamical variables, the canonical transformations tend to break down the differences in character between the quantized field variables and the classical space-time continuum. In contrast to Snyder's assumptions,[6] our coordinates commute with each other, but not with the energy-momentum densities. The dynamical character of any particle coordinates follows automatically, but probably does not exhaust the physical significance of the coordinate commutation relations.

Throughout the mathematical discussion of the earlier sections it has been assumed implicitly that the matrix Λ, (3.13), has only eight null vectors. This assumption will not hold if the theory is covariant with respect to further transformation groups. Such a group will be the gauge transformation group if the theory is to contain an electromagnetic field with a conservation law of charge. In the event of any such additional covariance properties, the number of algebraic constraints will be increased correspondingly. They may be added to the Hamiltonian, and they will, in any case, commute with the Hamiltonian. The basic structure of the theory is not affected by such additional constraints.

** If this coefficient were permitted to vanish, the resulting choice of the parameter system would establish a linear dependence between du^s and dt.

[5] P. Weiss, Proc. Roy. Soc. **A156**, 192 (1936); **A169**, 102, 119 (1938).

[6] H. Snyder, Phys. Rev. **71**, 38 (1947); **72**, 68 (1947).

Chapter 34
P. A. M. Dirac (1950): Generalized Hamiltonian Dynamics

P. A. M. Dirac (1950). Generalized Hamiltonian Dynamics. *Canadian Journal of Mathematics*, 2: 129–148.

GENERALIZED HAMILTONIAN DYNAMICS

P. A. M. DIRAC

1. Introduction. The equations of dynamics were put into a general form by Lagrange, who expressed them in terms of a set of generalized coordinates and velocities. An alternative general form was later given by Hamilton, in terms of coordinates and momenta. Let us consider the relative merits of the two forms.

With the Lagrangian form the requirements of special relativity can very easily be satisfied, simply by taking the action, i.e. the time integral of the Lagrangian, to be Lorentz invariant. There is no such simple way of making the Hamiltonian form relativistic.

For the purpose of setting up a quantum theory one must work from the Hamiltonian form. There are well-established rules for passing from Hamilton's dynamics to quantum dynamics, by making the coordinates and momenta into linear operators. The rules lead to definite results in simple cases and, although they cannot be applied to complicated examples without ambiguity, they have proved to be adequate for practical purposes.

Thus both forms have their special values at the present time and one must work with both. The two forms are closely connected. Starting with any Lagrangian one can introduce the momenta and, in the case when the momenta are independent functions of the velocities, one can obtain the Hamiltonian. The present paper is concerned with setting up a more general theory which can be applied also when the momenta are not independent functions of the velocities. A more general form of Hamiltonian dynamics is obtained, which can still be used for the purpose of quantization, and which turns out to be specially well suited for a relativistic description of dynamical processes.

2. Strong and weak equations. We consider a dynamical system of N degrees of freedom, described in terms of generalized coordinates $q_n (n = 1,2, \dots, N)$ and velocities dq_n/dt or $\dot{q}_n$. We assume a Lagrangian L, which for the present can be any function of the coordinates and velocities

$$L \equiv L(q, \dot{q}). \tag{1}$$

We define the momenta by

$$p_n = \partial L/\partial \dot{q}_n. \tag{2}$$

For the development of the theory we introduce a variation procedure, varying each of the quantities q_n, $\dot{q}_n$, p_n independently by a small quantity δq_n, $\delta \dot{q}_n$, δp_n of order ϵ and working to the accuracy of ϵ. As a result of this

This paper is based on the first half of a course of lectures given at the Canadian Mathematical Seminar in Vancouver in August 1949.

variation procedure equation (2) will get violated, as its left-hand side will be made to differ from its right-hand side by a quantity of order ϵ. We shall now have to distinguish between two kinds of equations, equations such as (2) which get violated by a quantity of order ϵ when we apply the variation, and equations which remain valid to the accuracy ϵ under the variation. Equation (1) will be of the latter kind, since the variation in L will equal, by definition, the variation of the function $L(q,\dot{q})$. The former kind of equation we shall call a *weak equation* and write with the usual equality sign $=$, the latter we shall call a *strong equation* and write with the sign $\equiv$.

We have the following rules governing algebraic work with weak and strong equations:

$$\text{if } \quad A \equiv 0 \quad \text{then} \quad \delta A = 0;$$
$$\text{if } \quad X = 0 \quad \text{then} \quad \delta X \neq 0;$$

in general. From the weak equation $X = 0$ we can deduce

$$\delta X^2 = 2X\delta X = 0,$$

so we can deduce the strong equation

$$X^2 \equiv 0.$$

Similarly, from two weak equations $X_1 = 0$ and $X_2 = 0$ we can deduce the strong equation

$$X_1 X_2 \equiv 0.$$

It may be that the N quantities $\partial L/\partial \dot{q}_n$ on the right-hand side of (2) are all independent functions of the N velocities $\dot{q}_n$. In this case equations (2) determine each $\dot{q}$ as a function of the q's and p's. This case will be referred to as the *standard case*, and is the only one usually considered in dynamical theory.

If the $\partial L/\partial \dot{q}$'s are not independent functions of the velocities, we can eliminate the $\dot{q}$'s from equations (2) and obtain one or more equations

$$\phi(q, p) = 0 \tag{3}$$

involving only q's and p's. We may suppose equation (3) to be written in such a way that the variation procedure changes ϕ by a quantity of order ϵ, since if it changes ϕ by a quantity of order ϵ^k, we have only to replace ϕ by $\phi^{1/k}$ in (3) and the desired condition will be fulfilled. We now have equation (3) violated by the order ϵ when we apply the variation, so it is correctly written as a weak equation.

We shall need to use a complete set of independent equations of the type (3), say

$$\phi_m(q, p) = 0, \qquad m = 1,2, \ldots, \text{M}. \tag{4}$$

The condition of independence means that none of the ϕ's is expressible linearly in terms of the others, with functions of the q's and p's as coefficients. The condition of completeness means that any function of the q's and p's which vanishes on account of equations (2) and changes by the order ϵ with the variation procedure is expressible as a linear function of the ϕ_m with functions of the q's and p's as coefficients.

We may picture the relationship of strong and weak equations in the following way. Take the 3N dimensional space with the q's, $\dot{q}$'s and p's as coordinates. In this space there will be a 2N dimensional region in which equations (2) are satisfied. Call it the region R. Equations (4) will also be satisfied in this region, as they are consequences of (2). Now consider all points of the 3N dimensional space which are within a distance of order ϵ from R. They will form a 3N dimensional region like a shell with a thickness of order ϵ. Call this the region R_ϵ. *A weak equation holds in the region R, a strong equation holds in the region R_ϵ.*

3. The Hamiltonian. The Hamiltonian H is defined by

$$H \equiv p_n \dot{q}_n - L, \tag{5}$$

where a summation is understood over all values for a repeated suffix in a term. We have

$$\begin{aligned} \delta H &= \delta(p_n \dot{q}_n - L) \\ &= p_n \delta\dot{q}_n + \dot{q}_n \delta p_n - \partial L/\partial q_n . \delta q_n - \partial L/\partial \dot{q}_n . \delta \dot{q}_n \\ &= \dot{q}_n \delta p_n - \partial L/\partial q_n . \delta q_n. \end{aligned} \tag{6}$$

We find that δH does not depend on the $\delta\dot{q}$'s. This important result holds whether we have the standard case or not.

Equation (5) gives a definition for H as a function of the q's, $\dot{q}$'s and p's, holding throughout the 3N dimensional space of q's, $\dot{q}$'s, and p's. We shall use the definition only in the region R_ϵ, and in this region the result (6) holds, to the first order. This means that, if we keep the q's and p's constant and make a first-order variation in the $\dot{q}$'s, the variation in H will be of the second order. Thus if we keep the q's and p's constant and make a finite variation in the $\dot{q}$'s, keeping all the time in the region R_ϵ (which is possible when we do not have the standard case), the variation in H will be of the first order. If we keep in the region R, the variation in H will be zero. It follows that in the region R, H is a function of the q's and p's only. Calling this function $\mathfrak{H}(q, p)$, we have the weak equation

$$H = \mathfrak{H}(q, p) \tag{7}$$

holding in the region R. In the standard case the function $\mathfrak{H}$ is the ordinary Hamiltonian.

Starting from a point in R and making a general variation, we have from

$$\delta(H - \mathfrak{H}) = \left(\dot{q}_n - \frac{\partial \mathfrak{H}}{\partial p_n}\right)\delta p_n - \left(\frac{\partial L}{\partial q_n} + \frac{\partial \mathfrak{H}}{\partial q_n}\right)\delta q_n. \tag{6}$$

Thus $\delta(H - \mathfrak{H})$ depends only on the δq's and δp's. If the variation is such that we stay in the region R, then of course $\delta(H - \mathfrak{H}) = 0$. Thus $\delta(H - \mathfrak{H})$ vanishes for any variation of the q's and p's such that one can choose the $\delta\dot{q}$'s so as to preserve equations (2). The only restriction this imposes on the δq's and δp's is that they must preserve equations (4), i.e. they must lead to $\delta\phi_m = 0$ for all m. Thus $\delta(H - \mathfrak{H})$ is zero for any values δq, δp that make $\delta\phi_m = 0$, and hence for arbitrary δq, δp

(8) $$\delta(H - \mathfrak{H}) = v_m \delta\phi_m$$

with suitable coefficients v_m. These coefficients will be functions of the q's, $\dot{q}$'s and p's, and with the help of (2) can be expressed as functions of the q's and $\dot{q}$'s only. We now get

$$\delta(H - \mathfrak{H} - v_m\phi_m) = \delta(H - \mathfrak{H}) - v_m\delta\phi_m - \phi_m\delta v_m = 0$$

from (8) and (4), and hence

(9) $$H \equiv \mathfrak{H} + v_m\phi_m.$$

We have here a strong equation, holding to the first order in the region R_ϵ, in contradistinction to the weak equation (7), which holds only in R.

Equation (8) gives

$$\delta H = \delta\mathfrak{H} + v_m\delta\phi_m$$

$$= \frac{\partial\mathfrak{H}}{\partial p_n}\,\delta p_n + \frac{\partial\mathfrak{H}}{\partial q_n}\,\delta q_n + v_m\left(\frac{\partial\phi_m}{\partial p_n}\,\delta p_n + \frac{\partial\phi_m}{\partial q_n}\,\delta q_n\right).$$

Comparing this with (6), we get

(10) $$\dot{q}_n = \frac{\partial\mathfrak{H}}{\partial p_n} + v_m\,\frac{\partial\phi_m}{\partial p_n},$$

(11) $$-\frac{\partial L}{\partial q_n} = \frac{\partial\mathfrak{H}}{\partial q_n} + v_m\,\frac{\partial\phi_m}{\partial q_n}.$$

Equations (10) give the $\dot{q}$'s in terms of the q's, p's and v's. They show that the 2N variables $q_n, \dot{q}_n$ can be expressed in terms of the 2N + M variables q_n, p_n, v_m. Between these 2N + M variables there exist the M relations (4). There cannot be any other relations between these variables, as otherwise the 2N variables q_n, $\dot{q}_n$ would not be independent. Thus the v's must each be independent of the q's, p's and other v's. The v's can be considered as a kind of velocity variables, which serve to fix those $\dot{q}$'s that cannot be expressed in terms of q's and p's.

When we work with the Hamiltonian form of dynamics we use as basic variables the q's, p's and v's, between which certain relations (4) are assumed to exist, and which are otherwise independent. These variables will be called the *Hamiltonian variables*.

4. The equations of motion. We assume the usual Lagrangian equations of motion as weak equations,

(12) $$\dot{p}_n = \partial L/\partial q_n.$$

By substituting for the p's in (12) their values given by (2), we get equations involving the accelerations $\ddot{q}_n$. In the standard case these equations will determine all the $\ddot{q}$'s in terms of the q's and $\dot{q}$'s. In the case with M equations (4), the equations of motion will give us only N − M equations for the $\ddot{q}$'s. The remaining M equations of motion will tell us how the ϕ_m's vary with time.

For consistency the ϕ_m's must remain zero. These consistency conditions will be examined later.

With the help of (11) the equations of motion (12) take the form

$$\dot{p}_n = -\frac{\partial \mathfrak{H}}{\partial q_n} - v_m \frac{\partial \phi_m}{\partial q_n}. \tag{13}$$

Equations (13) together with (10) constitute the Hamiltonian equations of motion. They are fixed by the function $\mathfrak{H}$ and the equations $\phi_m = 0$. The Hamiltonian equations of motion give us the $\dot{q}$'s and $\dot{p}$'s in terms of the Hamiltonian variables q, p, v. They give us no direct information about the $\dot{v}$'s, but will give us some information indirectly when we examine the consistency conditions.

The Hamiltonian equations of motion can be expressed more easily with the help of the Poisson bracket notation. Any two functions ξ and η of the q's and p's have a P. b. $[\xi, \eta]$, defined by

$$[\xi, \eta] \equiv \frac{\partial \xi}{\partial q_n}\frac{\partial \eta}{\partial p_n} - \frac{\partial \xi}{\partial p_n}\frac{\partial \eta}{\partial q_n}. \tag{14}$$

It is easily verified that the P. b. remains invariant under a transformation to new q's and p's, in which the new q's are any independent functions of the original q's and the new p's are defined by the new equations (2) with L expressed in terms of the new q's and their time derivatives. This invariance property gives the P. b. its importance.

P. b.'s are subject to the following laws, which are easily verified from the definition:

$$\begin{cases} [\xi, \eta] \equiv -[\eta, \xi], \\ [\xi, f(\eta_1, \eta_2, \ldots)] \equiv \dfrac{\partial f}{\partial \eta_1}[\xi, \eta_1] + \dfrac{\partial f}{\partial \eta_2}[\xi, \eta_2] + \ldots, \\ [\xi, [\eta, \zeta]] + [\eta, [\zeta, \xi]] + [\zeta, [\xi, \eta]] \equiv 0. \end{cases} \tag{15}$$

In the second of these laws f is any function of various quantities $\eta_1, \eta_2, \ldots$, each of which is a function of the q's and p's. The third law, known as Poisson's identity, applies to any three functions ξ, η, ζ of the q's and p's.

It is desirable to extend the notion of P. b.'s to include functions of the $\dot{q}$'s which are not expressible in terms of the q's and p's. We assume these more general P. b.'s are subject to the laws (15) but are otherwise arbitrary. Alternatively, we may assume that the $\dot{q}$'s are arbitrary functions of the q's and p's, and the laws (15) can then be deduced with ξ, the η's and ζ involving the $\dot{q}$'s.

From a strong equation $A \equiv 0$ we can infer the weak equations

$$\frac{\partial A}{\partial q_n} = 0, \quad \frac{\partial A}{\partial \dot{q}_n} = 0, \quad \frac{\partial A}{\partial p_n} = 0,$$

and hence, by an application of the second of the laws (15),

$$[\xi, A] = 0$$

for any ξ. We may have $[\xi, A] \equiv 0$, (for example when $A \equiv 0$ by definition) but this is not necessarily so. From a weak equation $X = 0$ we cannot infer $[\xi, X] = 0$ in general.

If g is any function of the q's and p's, we have from (10) and (13)

$$\dot{g} = \frac{\partial g}{\partial q_n}\left(\frac{\partial \mathfrak{H}}{\partial p_n} + v_m \frac{\partial \phi_m}{\partial p_n}\right) - \frac{\partial g}{\partial p_n}\left(\frac{\partial \mathfrak{H}}{\partial q_n} + v_m \frac{\partial \phi_m}{\partial q_n}\right)$$

$$= [g, \mathfrak{H}] + v_m [g, \phi_m]. \tag{16}$$

This is the general Hamiltonian equation of motion. It may also be written, with the help of (4), as

$$\dot{g} = [g, \mathfrak{H}] + v_m [g, \phi_m] + [g, v_m]\, \phi_m = [g, H], \tag{17}$$

when it is the same as the usual Hamiltonian equation of motion in P. b. notation.

5. Homogeneous velocities. The theory takes a specially simple form in the case when the Lagrangian is homogeneous of the first degree in the velocities. The momenta defined by (2) are then homogeneous of degree zero in the $\dot{q}$'s and so depend only on the ratios of the $\dot{q}$'s. Since there are N p's and only N − 1 independent ratios of the $\dot{q}$'s, the p's now cannot be independent functions of the $\dot{q}$'s and there must be at least one relation (4) connecting the q's and p's. The case when there is only one relation between the q's and p's may now be considered as the standard case.

From Euler's theorem we have

$$L \equiv \dot{q}_n\, \partial L/\partial \dot{q}_n \tag{18}$$

and hence

$$L = \dot{q}_n p_n$$

so that

$$H = 0. \tag{19}$$

This weak equation holding in the region R allows us to take $\mathfrak{H} \equiv 0$, so that (9) becomes

$$H \equiv v_m \phi_m. \tag{20}$$

The general equation of motion (16) is now

$$\dot{g} = v_m [g, \phi_m]. \tag{21}$$

The Hamiltonian equations of motion are now fixed entirely by the equations $\phi_m = 0$.

Equation (21) is homogeneous in the v's on the right-hand side. Given any solution of the equations of motion, one can obtain another solution from it by multiplying all the v's by a factor γ, which may vary arbitrarily with the time. The new solution will have the time rate of change of all dynamical variables multiplied by the factor γ. The new solution would be obtained from the previous solution if we replaced the time t by a new independent variable τ such that $dt/d\tau = \gamma$. The new independent variable is completely arbitrary: it can be any function of t and the q's and $\dot{q}$'s. Thus, given any solution of the equations of motion, we can get another solution from it by replacing t by an arbitrary τ, so *the equations of motion give us no information about the independent variable.* This is an important feature of dynamical theory with homogeneous velocities, and makes it specially convenient for a relativistic treatment.

The Lagrangian for any dynamical system can be made to satisfy the condition for homogeneous velocities by taking the time t to be an extra coordinate q_0 and using the equation $\dot{q}_0 = 1$ to make the Lagrangian homogeneous of the first degree in all the velocities, including $\dot{q}_0$. The new Lagrangian equations of motion for all the q's can then be deduced, as has been shown by the author [1]. In this way we can get a new formulation for a general dynamical system in terms of homogeneous velocities. The new formulation gives all the equations of the old formulation except the equation $\dot{q}_0 = 1$. If we want to have this equation in the new formulation we may assume it as a supplementary condition, not derivable from the equations of motion but consistent with them. We can, however, very well dispense with it, as its only effect is to fix the independent variable, which would otherwise be arbitrary in the homogeneous velocity formulation.

Thus we may confine ourselves to the homogeneous velocity theory without losing any generality. We shall do this in future as it leads to somewhat simpler equations, and use the dot to denote differentiation with respect to an arbitrary independent variable τ.

6. The consistency conditions. For consistency the equations of motion must make each of the ϕ_m remain zero. Thus, putting $\phi_{m'}$ for g in (21), we get

$$v_m [\phi_m, \phi_{m'}] = 0. \tag{22}$$

Let us suppose the equations (22) to be reduced as far as possible with the help of the set of equations (4). The reduction may involve the cancellation of factors when we can assume these factors do not vanish. The resulting equations must each be of one of four types.

Type 1. It involves some of the variables v.

Type 2. It is independent of the v's but involves some of the variables p and q. It is thus of the form

$$\chi(q, p) = 0 \tag{23}$$

and is independent of the equations (4).

Type 3. It reduces to $0 = 0$.

Type 4. It reduces to $1 = 0$.

An equation of type 2 leads to a further consistency condition, since we must have χ remaining zero. Putting χ for g in (21), we get

$$v_m[\phi_m, \chi] = 0. \tag{24}$$

This equation, reduced as far as possible with the help of equations (4) and any equations (23) that we already have, will again be of one of the four types.

If it is of type 2 it will lead to yet another consistency condition. We continue in this way with each equation of type 2 until it leads to an equation of another type.

If any of the equations obtained in this way is of type 4, the equations of motion are inconsistent. This case is of no interest and will be excluded in future. Equations of type 3 are automatically satisfied. We are left with the equations of types 1 and 2 to fit into the theory.

Let us call the complete set of equations of type 2

$$\chi_k(q,p) = 0, \qquad k = 1, 2, \ldots, \text{K}. \tag{25}$$

We may suppose the functions χ_k to be chosen, like the ϕ_m in (4), so that their variations are of order ϵ. Equations (25) are then correctly written as weak equations. These further weak equations will reduce the region R, in which all weak equations hold, so as to have only $2\text{N} - \text{K}$ dimensions. The region R_ϵ will also get reduced, as it will now consist of all points within a distance of order ϵ from the new region R.

For studying the equations of type 1, it is convenient to introduce some new concepts. We define one of the quantities ϕ_m to be a *first class* ϕ if its P. b. with every ϕ and χ vanishes. Thus $\phi_{m'}$ is first class if

$$\begin{aligned} [\phi_{m'}, \phi_m] &= 0, \qquad & m &= 1, 2, \ldots, \text{M}, \\ [\phi_{m'}, \chi_k] &= 0, \qquad & k &= 1, 2, \ldots, \text{K}. \end{aligned} \tag{26}$$

These equations need hold only in the weak sense, which means that they need hold only as consequences of the equations $\phi_m = 0$, $\chi_k = 0$. Thus the left-hand sides of (26) must each equal in the strong sense some linear function of the ϕ_m and χ_k. A ϕ which does not satisfy all these conditions we call a *second class* ϕ.

We can make a linear transformation of the ϕ's of the form

$$\phi^*_m = \gamma_{mm'}\phi_{m'}, \tag{27}$$

where the γ's are any functions of the q's and p's such that their determinant does not vanish in the weak sense. The ϕ^*'s are then equivalent to the ϕ's for all the purposes of the theory.

Let us make a transformation of this kind so as to bring as many ϕ's as possible into the first class. Let us call the first class ϕ's that we then have ϕ_α's and the second class ones ϕ_β's, with $\beta = 1, 2, \ldots \text{B}$ and $\alpha = \text{B} + 1$, $\text{B} + 2, \ldots \text{M}$.

If $\phi_{m'}$ is first class, equation (22) is automatically satisfied. Further, in equations (22) and (24) we can restrict ϕ_m to be second class, as first class ϕ_m's contribute zero. Thus the surviving equations (22) and (24) will read

$$\begin{aligned} v_\beta[\phi_\beta, \phi_{\beta'}] &= 0, \qquad & \beta, \beta' &= 1, 2, \ldots, \text{B}, \\ v_\beta[\phi_\beta, \chi_k] &= 0, \qquad & k &= 1, 2, \ldots, \text{K}. \end{aligned} \tag{28}$$

These are all the equations of type 1. They show that *either* all the v_β's vanish *or* the matrix

$$(29)\qquad \left\| \begin{matrix} 0 & [\phi_1, \phi_2] & [\phi_1, \phi_3] & \cdots & [\phi_1, \phi_B] & [\phi_1, \chi_1] & \cdots & [\phi_1, \chi_K] \\ [\phi_2,\phi_1] & 0 & [\phi_2, \phi_3] & \cdots & [\phi_2,\phi_B] & [\phi_2, \chi_1] & \cdots & [\phi_2, \chi_K] \\ \cdot & \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ [\phi_B,\phi_1] & [\phi_B, \phi_2] & [\phi_B,\phi_3] & \cdots & 0 & [\phi_B,\chi_1] & \cdots & [\phi_B, \chi_K] \end{matrix} \right\|$$

is of rank less than B, in the weak sense.

It will now be proved that the first alternative is the correct one. Assume the matrix (29) is of rank U $<$ B. Form the determinant

$$(30)\qquad D \equiv \begin{vmatrix} \phi_1 & 0 & [\phi_1,\phi_2] & [\phi_1,\phi_3] & \cdots & [\phi_1,\phi_U] \\ \phi_2 & [\phi_2,\phi_1] & 0 & [\phi_2,\phi_3] & \cdots & [\phi_2,\phi_U] \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot \\ \phi_{U+1} & [\phi_{U+1},\phi_1] & [\phi_{U+1},\phi_2] & [\phi_{U+1},\phi_3] & \cdots & [\phi_{U+1},\phi_U] \end{vmatrix}.$$

D is a linear function of the ϕ_β's and so vanishes in the weak sense. The P. b. of D with any quantity f equals the sum of the determinants formed by taking the P. b. of each column of (30) with f. All these determinants, except the one formed by taking the P. b. of the first column with f, will vanish in the weak sense, as the elements of their first column all vanish in the weak sense. Thus

$$(31)\qquad [D, f] = \begin{vmatrix} [\phi_1,f] & 0 & [\phi_1,\phi_2] & [\phi_1,\phi_3] & \cdots & [\phi_1,\phi_U] \\ [\phi_2,f] & [\phi_2,\phi_1] & 0 & [\phi_2,\phi_3] & \cdots & [\phi_2,\phi_U] \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot \\ [\phi_{U+1},f] & [\phi_{U+1},\phi_1] & [\phi_{U+1},\phi_2] & [\phi_{U+1},\phi_3] & \cdots & [\phi_{U+1},\phi_U] \end{vmatrix}.$$

If we take f to be any ϕ_a, the first column of (31) vanishes and so $[D, \phi_a] = 0$. If we take f to be any ϕ_β or χ, the determinant (31) either has two columns identical and so vanishes, or it is a minor of the matrix (29) with U + 1 rows and columns, and vanishes because this matrix is assumed to be of rank U. Thus D has zero P. b. with all the ϕ's and χ's.

It may be that D vanishes in the strong sense on account of the co-factors of the elements of its first column all vanishing in the weak sense. If this is the case, we take a different determinant D, with its columns after the first one corresponding to any U of the columns of (29) and its rows corresponding to any U + 1 of the rows of (29). We can always choose such a determinant D so that the co-factors of the elements of its first column do not all vanish, from the assumption that (29) is of rank U. We get in this way a D which is a first class ϕ and is a linear function of the ϕ_β's. This contradicts the assumption that we had previously put as many ϕ's as possible in the first class.

We can conclude that *if we have put as many ϕ's as possible in the first class, the v's associated with the second class ϕ's all vanish.* The Hamiltonian (20) then reduces to

$$(32)\qquad H \equiv v_a\phi_a,$$

and the general equation of motion (21) becomes

$$(33)\qquad \dot{g} = v_a[g, \phi_a].$$

The vanishing of the v_β's together with the equations (25) ensure that the consistency conditions are all satisfied. The v_a's remain completely arbitrary. Each of them gives rise to a freedom of motion—an arbitrary function in the general solution of the equations of motion. In the standard case there is just one ϕ, which is necessarily first class, and thus there is one arbitrary function in the general solution of the equations of motion. This is connected with the arbitrary character of the independent variable τ.

7. Supplementary conditions. In dealing with a particular dynamical system, we may wish to impose equations on the coordinates and velocities additional to the equations of motion that follow from the Lagrangian. Such supplementary conditions must be introduced as further weak equations in the theory.

With the help of equations (10) (with $\mathfrak{H} \equiv 0$) the supplementary conditions can be expressed as relations between the q's, p's and v's. They may lead to equations between the q's and p's only. Such equations must be treated as extra χ equations, to be joined on to the set (25). They will give rise to further consistency conditions, which are to be handled in the same way as the preceding ones and may lead to still more χ equations. A first class ϕ must now be defined to have zero P. b. also with these new χ's, so the number of first class ϕ's may be reduced by the supplementary conditions. This would cause a reduction in the number of freedoms of motion.

Those of the supplementary conditions that do not give χ equations will give conditions on the v_a variables. These conditions will usually be of a more complicated kind than merely the vanishing of certain v's, like all the conditions on the v's that follow from consistency conditions. They will make a further reduction in the number of freedoms of motion, reducing it to less than the number of first class ϕ's.

8. Transformations of the Hamiltonian form. Take a set of functions $\theta_s(s = 1, 2, \ldots \mathrm{s})$ of the q's and p's such that the determinant

$$\Delta \equiv \begin{vmatrix} 0 & [\theta_1, \theta_2] & [\theta_1, \theta_3] & \cdots & [\theta_1, \theta_\mathrm{s}] \\ [\theta_2, \theta_1] & 0 & [\theta_2, \theta_3] & \cdots & [\theta_2, \theta_\mathrm{s}] \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ [\theta_\mathrm{s}, \theta_1] & [\theta_\mathrm{s}, \theta_2] & [\theta_\mathrm{s}, \theta_3] & \cdots & 0 \end{vmatrix} \tag{34}$$

does not vanish in the weak sense. This implies that s must be even. Let $c_{ss'}$ denote the co-factor of $[\theta_s, \theta_{s'}]$ divided by Δ, so that

$$c_{ss'} \equiv - c_{s's}$$

and

$$c_{ss'}[\theta_s, \theta_{s''}] \equiv \delta_{s's''} \ . \tag{35}$$

Then we can define a new P. b. $[\xi, \eta]^*$ for any two quantities ξ and η by the formula

$$[\xi, \eta]^* \equiv [\xi, \eta] + [\xi, \theta_s]\, c_{ss'}\, [\theta_{s'}, \eta]. \tag{36}$$

It is easily seen that the new P. b.'s obey the first two of the laws (15), and after some calculation one finds that they also obey the third, Poisson's identity. (See Appendix.) The new P. b.'s make

$$\begin{aligned}[\xi, \theta_s]^* &\equiv [\xi, \theta_s] + [\xi, \theta_{s'}]\, c_{s's''}[\theta_{s''}, \theta_s] \\ &\equiv [\xi, \theta_s] - [\xi, \theta_{s'}]\, \delta_{s's} \\ &\equiv 0 \qquad (37)\end{aligned}$$

for any ξ.

To understand the significance of the new P. b.'s, let us take the case when the θ's consist of $\frac{1}{2}$s coordinates q and their conjugate p's. We then see that the new P. b.'s are obtained by omitting the terms involving differentiations with respect to these q's and p's from the summation over n in the definition (14). Thus the new P. b.'s refer to a system of N $- \frac{1}{2}$s degrees of freedom. If, instead of taking the θ's to be just certain q's and p's, we take them to be any independent functions of these q's and p's, we get the same new P. b.'s. With general θ's the new P. b.'s will still refer to a system of N $- \frac{1}{2}$s degrees of freedom, but the reduction of the degrees of freedom is made in a more complicated way than the mere omission of certain q's and p's.

Let us suppose the θ's are all ϕ's or χ's. (The ϕ's must be second class, as otherwise $\Delta = 0$.) We then have $[\theta_s, H] = 0$ for all s, and hence

$$[g, H]^* = [g, H] = \dot{g} \qquad (38)$$

for g any function of the q's and p's. Thus *the new P. b.'s may be used to give the Hamiltonian equations of motion.* We get in this way a new form for the equations of motion, which is simpler because the number of effective degrees of freedom is reduced.

Each of the θ's now vanishes in the weak sense. If we work only with the new P. b.'s, we can assume each of the θ's vanishes in the strong sense without getting a contradiction, because from (37) the new P. b. of a θ with anything vanishes. We can then use the strong equations $\theta_s \equiv 0$ to simplify the Hamiltonian.

Let us define a χ to be first class it if has zero P. b. with all the ϕ's and χ's and to be second class otherwise. We can make a linear transformation of the χ's of the form

$$\chi^*_k = \gamma_{kk'}\chi_{k'} + \gamma'_{km}\,\phi_m, \qquad (39)$$

where the γ's and γ''s are any functions of the q's and p's such that the determinant of the γ's does not vanish in the weak sense, and the new χ's are then equivalent to the old ones for all the purposes of the theory. Let us make a transformation of this kind so as to bring as many χ's as possible into the first class, and let us call the first class χ's that we then have χ_a's and the second class ones χ_β's.

We may take the θ's to consist of all the ϕ_β's and χ_β's. The determinant Δ then does not vanish. The proof of this result is similar to the proof that the matrix (29) is of rank B, and consists in assuming that Δ is of rank T $<$ S and constructing a determinant like

$$(40)\qquad \begin{vmatrix} \theta_1 & 0 & [\theta_1, \theta_2] & \cdots & [\theta_1, \theta_T] \\ \theta_2 & [\theta_2, \theta_1] & 0 & \cdots & [\theta_2, \theta_T] \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \theta_{T+1} & [\theta_{T+1}, \theta_1] & [\theta_{T+1}, \theta_2] & \cdots & [\theta_{T+1}, \theta_T] \end{vmatrix},$$

which is then seen to be a first class ϕ or χ and is a linear function of the ϕ_β's and χ_β's, so it contradicts the assumption that as many ϕ's and χ's as possible have been put in the first class.

With this choice of θ's we get the maximum simplification of the Hamiltonian equations of motion by this method. We get a new scheme in which all the ϕ_β and χ_β equations are strong equations. We may be able to use these equations to eliminate some of the q's and p's entirely from the theory.

The form of the new scheme is not unique, because the ϕ_β's and χ_β's are not unique. If we merely replace the ϕ_β's and χ_β's by linear functions of themselves, we do not change the final form. We can, however, add to the ϕ_β's any linear functions of the ϕ_a's, and to the χ_β's any linear functions of the ϕ_a's and χ_a's, which does not change Δ or the $c_{ss'}$, but does in general change $[\xi, \eta]^*$, and so the form of the Hamiltonian scheme is altered. The different forms must, of course, be equivalent, as they all give the same equations of motion.

As an application of the above method, let us consider the case of a Lagrangian that does not involve some of the velocities. Suppose L does not involve $\dot{q}_j$ ($j = 1, 2, \ldots, J < N$). Then each p_j equals zero in the weak sense and equals a ϕ in the strong sense. Suppose that no linear combination of the p_j's is first class. Then we can take the p_j's to be ϕ_β's. Let us now take half the θ's to be the p_j's and the other half to be suitable second class ϕ's or χ's so that Δ does not vanish. Call these other θ's θ_j. With this choice of θ's one easily sees that the new P. b.'s are just what one would get if one applied the definition (14) to those degrees of freedom for which q is not a q_j, with each p_j reckoned as strongly equal to zero and each q_j reckoned as strongly equal to a function of the other q's and p's given by the equations $\theta_j \equiv 0$. We get in this way a new Hamiltonian scheme (not necessarily with the maximum simplification, as there may be other ϕ_β's and χ_β's not included in the θ's) in which the q_j and p_j do not appear as independent dynamical variables.

The new scheme could be obtained in a more direct way by not counting the q_j's as coordinates right from the beginning, and not introducing momenta conjugate to them at all. Let us see what modifications this would bring into the development of the theory.

Define n so as to take on only those values for which q is not a q_j, i.e. the values $J+1, J+2, \ldots, N$. Then equations (2) and (5) still hold and equation (6) must be replaced by

$$(41)\qquad \delta H = \dot{q}_n \delta p_n - \partial L/\partial q_n . \delta q_n - \partial L/\partial q_j . \delta q_j,$$

as we allow the q_j's to vary. We may assume the equations

$$(42)\qquad \partial L/\partial q_j = 0$$

as supplementary conditions with this method. Equation (41) then reduces to precisely (6). We can infer that H is of the form (20), where the ϕ_m are functions of the q_n's and p_n's, independent of the q_j's, that vanish on account of equations (2). The remainder of the theory can be developed as before, in terms of ϕ's and χ's that do not involve the q_j's . Those of the ϕ or χ equations that do involve the q_j's can be looked upon as defining the q_j's in terms of the other variables, and play no further role in the theory.

With this form of the theory we have the Lagrangian containing variables q_j that involve momenta. The appearance of momentum variables in the Lagrangian is analogous to the appearance of the velocity variables v_a in the Hamiltonian.

9. The Hamiltonian as starting point. Instead of starting with the Lagrangian and obtaining the Hamiltonian from it, one can start with the Hamiltonian. One begins by assuming certain dynamical variables q_n and p_n $(n = 1, 2, \ldots, \text{N})$, or maybe other dynamical variables between which there are definite P. b. relations satisfying the laws (15), and assuming certain weak equations as ϕ equations connecting them. There is no point in distinguishing ϕ's and χ's with this method. At least one of the ϕ's must be first class, i.e. must have zero P. b. with all the ϕ's, or there can be no consistent motion. One then assumes the Hamiltonian to be a linear function of the first class ϕ's ϕ_a, with new variables v_a as coefficients, and assumes the Hamiltonian equations of motion (17) or (33). The v's can vary arbitrarily with the independent variable τ.

The previous scheme of equations of motion, derived from a Lagrangian and involving possibly χ's as well as ϕ's, is to be looked upon as an example of the present scheme with some of the v's restricted to be zero by supplementary conditions. The ϕ_a's corresponding to these v_a's are then the first class χ's of the previous scheme. Such supplementary conditions, or any supplementary conditions involving the v's, are of no value for the application of the theory to relativistic dynamics given in the next section and cannot be taken over into the quantum theory, so they will not be included in the further work. Supplementary conditions not involving the v's are just ϕ equations.

The P. b. of two first class ϕ's is a first class ϕ, as may be verified in the following way. The P. b. $[\phi_a, \phi_{a'}]$ vanishes weakly and so is strongly equal to a linear function of the ϕ's, these being the only quantities that are weakly zero in the present scheme. We have to show that its P. b. with an arbitrary ϕ is weakly zero. From Poisson's identity

$$[\phi, [\phi_a, \phi_{a'}]] \equiv [[\phi, \phi_a], \phi_{a'}] - [[\phi, \phi_{a'}], \phi_a]. \tag{43}$$

Since ϕ_a is first class, $[\phi, \phi_a]$ vanishes weakly and so is strongly equal to a linear function of the ϕ's, and hence its P. b. with the first class $\phi_{a'}$ vanishes weakly. Similarly the second term on the right-hand side of (43) vanishes weakly, so the required result is proved.

Suppose there are A independent first class ϕ's and M independent ϕ's altogether. In the phase space (the 2N-dimensional space of the q_n and p_n variables) there is a space of (2N − M) dimensions in which all the ϕ equations are satisfied. Call it the $(2N - M)$-space. The state of the dynamical system for a particular τ value is fixed by giving values to the q's and p's satisfying all the ϕ equations and is thus represented by a point P in the $(2N-M)$-space. The motion of the system ensuing from this state is represented by a curve in the $(2N - M)$-space starting from P. On account of the A variables v_a being arbitrary, this curve may start out in any direction in a small space of A dimensions surrounding P. There is one of these small spaces of A dimensions surrounding every point of the $(2N-M)$-space. It will now be shown that *these small spaces are integrable.*

Suppose that for an interval of τ, $\delta\tau = \epsilon_1$, all the v's vanish except $v_{a'}$, which is equal to 1, and that for the following τ interval, $\delta\tau = \epsilon_2$, all the v's vanish except $v_{a''}$, which is equal to 1. Then any function g of the q's and p's is changed at the end of the first interval to

$$g + \epsilon_1[g, \phi_{a'}].$$

It is changed at the end of the second interval, with the accuracy $\epsilon_1\epsilon_2$ but with neglect of ϵ_1^2 and ϵ_2^2, to

$$g + \epsilon_1[g, \phi_{a'}] + \epsilon_2[g + \epsilon_1[g, \phi_{a'}], \phi_{a''}]. \tag{44}$$

If the two kinds of motion are made in the reverse order, g changes to

$$g + \epsilon_2[g, \phi_{a''}] + \epsilon_1[g + \epsilon_2[g, \phi_{a''}], \phi_{a'}]. \tag{45}$$

The difference between (44) and (45) is, by Poisson's identity,

$$\epsilon_1\epsilon_2[g,[\phi_{a'}, \phi_{a''}]]. \tag{46}$$

It was shown above that $[\phi_{a'}, \phi_{a''}]$ is a first class ϕ, so that (46) is a possible change in g arising from the equations of motion with a suitable choice of the v's, and thus corresponds to a motion in the small A-dimensional space round the starting point. This is the condition for integrability.

If there are supplementary conditions involving the v's, this integrability may get spoiled. Thus the integrability does not necessarily hold for the equations of motion derived from a Lagrangian.

The integration of the small spaces will provide a set of A-dimensional spaces lying in the $(2N-M)$-space such that the motion always takes place entirely in one of them. Call these spaces A-spaces. Every curve in an A-space represents a possible solution of the equations of motion. Every point of the $(2N-M)$-space lies in an A-space, which contains all the motions starting from that point. It would be permissible to consider the A-space itself as the complete solution of the equations of motion, rather than a general curve in it.

Given a particular A-space, we can fix a point of it by A coordinates, each of which is some function of the q's and p's. Call these coordinates t_a ($a = 1, 2, \ldots$, A). They will play the role of time variables. The A-space itself can be described by giving all the q's and p's as functions of the t_a. If g is any q or p, or a function of the q's and p's, we have

$$\dot{g} = \dot{t}_a \, \partial g/\partial t_a, \tag{47}$$

since the τ variation of g may be looked upon as arising from the τ variation of the t_a's. Using the Hamiltonian equations of motion (33) for $\dot{g}$ and $\dot{t}_a$, we get

$$v_a[g, \phi_a] = v_a[t_a, \phi_a] \, \partial g/\partial t_a.$$

This equation holds for arbitrary v_a, so

$$[g, \phi_a] = [t_a, \phi_a] \, \partial g/\partial t_a. \tag{48}$$

Equations (48) may be looked upon as the general equations of motion that fix an A-space. They are the closest equations in the theory with homogeneous velocities to the usual Hamiltonian equations of motion. If $\mathrm{A} = 1$, we may take the one variable t_a to be the time and (48) then reduces to precisely the usual Hamiltonian equations of motion.

To pass from the Hamiltonian to the Lagrangian, we introduce the velocities $\dot{q}_n$ by the equations

$$\dot{q}_n = v_a \, \partial \phi_a/\partial p_n, \tag{49}$$

and then define L by

$$L \equiv p_n \dot{q}_n - H \equiv p_n \dot{q}_n - v_a \phi_a. \tag{50}$$

This gives L as a function of the q's, $\dot{q}$'s, p's and v's, linear in the $\dot{q}$'s and v's. Making independent variations δq, $\delta \dot{q}$, δp, δv, we get

$$\begin{aligned} \delta L &= \dot{q}_n \delta p_n + p_n \delta \dot{q}_n - \phi_a \delta v_a - v_a(\partial \phi_a/\partial q_n . \delta q_n + \partial \phi_a/\partial p_n . \delta p_n) \\ &= p_n \delta \dot{q}_n - v_a \, \partial \phi_a/\partial q_n . \delta q_n. \end{aligned} \tag{51}$$

Thus δL does not depend on the δp_n and δv_a. This result is to be compared with (6).

If the equations (49) together with the ϕ equations give the $\dot{q}$'s as independent functions of the p's and v's, so that they allow the p's and v's to be considered as functions of the q's and $\dot{q}$'s, then (51) shows that L is strongly equal to a function of the q's and $\dot{q}$'s only. This function must be homogeneous of the first degree in the $\dot{q}$'s. Differentiating it partially with respect to a q or $\dot{q}$, we find

$$\begin{aligned} \partial L/\partial \dot{q}_n &= p_n \\ \partial L/\partial q_n &= - v_a \, \partial \phi_a/\partial q_n = \dot{p}_n. \end{aligned} \tag{52}$$

These are the usual Lagrangian equations.

If the equations (49) together with the ϕ equations do not give the $\dot{q}$'s as independent functions of the p's and v's, they lead to certain equations between the q's and $\dot{q}$'s only, say

$$R_j(q, \dot{q}) = 0, \qquad j = 1, 2, \ldots, \mathrm{J}. \tag{53}$$

The R's are homogeneous in the $\dot{q}$'s and we arrange them to be of the first degree. We now proceed by a method analogous to that of §3 with the role of p's and $\dot{q}$'s interchanged. We obtain a result analogous to (9),

$$L \equiv \mathfrak{L} + \lambda_j R_j, \tag{54}$$

where $\mathfrak{L}$ is a function of the q's and $\dot{q}$'s only, which must be homogeneous of the first degree in the $\dot{q}$'s, and the coefficients λ_j are functions of the q's, p's and v's.

We have again equations (52) if the λ's are counted as independent variables in the partial differentiations of L, and L is then homogeneous of the first degree in the $\dot{q}$'s. Thus we have a Lagrangian containing momentum variables of the type considered at the end of the preceding section, with the previous q_j corresponding to the present λ_j and the supplementary conditions (42) giving the equations (53).

10. Application to relativistic dynamics. In the ordinary non-relativistic dynamics one works with the state of a dynamical system at a particular instant of time, this state being specified by giving values to the q's and p's. One has equations of motion which enable one, given the state at one instant, to calculate the state at another instant. These equations of motion, written in the Hamiltonian form with homogeneous velocities, need only one first class ϕ.

To get a dynamical theory which satisfies restricted relativity, we must set up a scheme of equations which applies equally to observers with all velocities. If we work with instants, we must include instants with respect to all observers. An instant is then any flat three-dimensional surface in space-time having a normal in a direction within the light-cone. A general instant needs four parameters to describe it, three to fix the direction of the normal, or the velocity of the observer, and the fourth to distinguish different instants for the same observer.

A relativistic dynamics that involves instants must enable one, given the state at any of these instants, to calculate the state at any other. We must have equations of motion showing how the dynamical variables vary as the instant varies. We can allow the instant to vary arbitrarily, with a translational motion in space-time as well as the direction of its normal varying, and the equations of motion must always apply. Thus *we need four first class ϕ's to give rise to the four freedoms of motion of the instant.*

The four parameters that describe the instant are to be treated as q's, subject to the equations of motion (17) or (33) along with the other q's and p's. They are distinguished from the other q's and p's in that it is specially convenient to take them as the t variables of the equations of motion (48). These equations then show directly how any q or p varies for a given variation of the instant.

There are other forms of relativistic dynamics not involving instants, which have been discussed by the author [2]. There is the point form, in which a state is defined with reference to a point in space-time. This form also needs four first class ϕ's, corresponding to the four freedoms of motion of the point. Then there is the front form, which needs three first class ϕ's, corresponding to the three freedoms of motion of a front. Finally, we may take a state to be defined on a general three-dimensional space-like surface in space-time. There must then be infinitely many first class ϕ's, corresponding to all the deformations that may be made in such a surface. With each of these forms the variables that describe the point, front, or general space-like surface are to be treated as q's, subject to the equations of motion (17) or (33), and are specially convenient to be taken as the t variables of equations (48).

The first class ϕ's discussed above are the fewest with which one can construct a relativistic dynamics in the respective forms. There may be additional ones. For example, an electrodynamics which allows gauge transformations to be made after one has fixed the initial values of all the q's and p's must contain extra freedoms of motion, which will need extra first class ϕ's to give rise to them.

11. Quantization. In order to quantize a dynamical system which one has worked out in the classical theory, one must set up a scheme of linear operators corresponding to the classical dynamical variables q and p, and to functions of them. There are no operators corresponding to the classical variables v, or to velocity variables in general, or to anything involving τ. The operators all operate on the vectors ψ of a Hilbert space, whose representatives in any representation are the wave functions which specify states in the quantum theory. Real classical variables correspond to self-adjoint operators.

The linear operators must be analogous to their classical counterparts in accordance with two general principles. Using the same letter to denote two things that are counterparts, the principles are

(i) P. b. relations between the classical variables correspond to commutation relations between the operators, according to the formula

$$[\xi, \eta] \quad \text{corresponds to} \quad 2\pi(\xi\eta - \eta\xi)/ih.$$

(ii) Weak equations between the classical variables correspond to linear conditions on the vectors ψ, according to the formula

$$X(q, p) = 0 \quad \text{corresponds to} \quad X\psi = 0.$$

The procedure of passing from the classical to the quantum theory is not mathematically well-defined, because whenever a classical quantity involves a product of two factors whose P. b. does not vanish, there is an ambiguity in the order in which the two factors should appear in the corresponding quantum expression. In practice with simple examples one finds no difficulty in deciding what the order should be. With complicated examples it may be impossible to choose the order in each case so as to make all the quantum equations consistent, and then one would not know how to quantize the theory. The present-day methods of quantization are all of the nature of practical rules, whose application depends on considerations of simplicity.

There are certain general features of the passage to the quantum theory that one must pay attention to, in order that the consistency of the quantum equations shall not go wrong in an elementary way. We have in the classical theory a number of ϕ equations (counting χ equations also as ϕ equations), which are to be used in the quantum theory according to the principle (ii). We can transform the classical ϕ's linearly by the transformation (27) and the new ϕ's are just as good as the old ones. If we make the corresponding transformation in the quantum theory, we must take care to put the coefficients γ all to the left of the ϕ's. *A general ϕ in the quantum theory is a linear function of the given ϕ's with coefficients on the left.*

From two quantum equations obtained from ϕ equations by principle (ii)

$$\phi_1\psi = 0, \qquad \phi_2\psi = 0,$$

we can infer

$$\phi_2\phi_1\psi = 0, \qquad \phi_1\phi_2\psi = 0,$$

and hence from principle (i),

$$[\phi_1, \phi_2,]\,\psi = 0.$$

This corresponds to the classical weak equation

$$[\phi_1, \phi_2] = 0.$$

We can infer that *all the ϕ's must be first class if the passage to the quantum theory is possible.*

Given a classical theory with second class ϕ's, one can get a quantum theory from it by first applying a transformation of the type described in §**8**, which converts all the ϕ_β equations into strong equations. The strong equations will correspond in the quantum theory to equations between operators, which serve to define some of them in terms of the others.

The quantum equations $\phi\psi = 0$, obtained by applying principle (ii) to the first class ϕ equations of the classical theory, are the Schroedinger wave equations. The usual classical dynamics with only one first class ϕ leads to only one Schroedinger equation. In the general theory there is one Schroedinger equation for each classical freedom of motion. The operators in these equations all correspond to classical dynamical variables for one τ value. The operators referring to a different τ value do not belong to the same algebraic scheme, and there does not seem to be anything in the quantum theory analogous to the τ dependence of the classical variables.

However, the dependence of the classical variables on the parameters t, given by equations (48), does have a quantum analogue, provided the t's are chosen so as to have zero P. b.'s with one another, so that they can be given numerical values simultaneously in the quantum theory. The specially convenient t's for the various forms of relativistic dynamics discussed in the preceding section do satisfy this condition. We cannot immediately take over equations (48) into the quantum theory because, as easily verified, the equations that we should get would not be invariant under a general linear transformation of the ϕ's (27). We must first put equations (48) in a standard form. By a transformation (27) we introduce a new set of ϕ's, ϕ_a say, in one-one correspondence with the t_a's, so that

$$[t_a, \phi_{a'}] = \delta_{aa'}. \tag{55}$$

With these ϕ's equations (48) reduce to

$$[g, \phi_a] = \partial g/\partial t_a. \tag{56}$$

These equations can be taken over into the quantum theory, and are then Heisenberg's quantum equations of motion for the present generalized dynamics.

Appendix. Proof of Poisson's identity for the new P. b.'s defined by (36).

Use the suffixes $r, s, t, \ldots$ to distinguish different θ's. We have by the definition

$$\begin{aligned}
&[[\xi,\eta]^*, \zeta]^* \\
&\quad = [[\xi,\eta] + [\xi,\theta_r]c_{rs}[\theta_s,\eta],\zeta] + [[\xi,\eta] + [\xi,\theta_r]c_{rs}[\theta_s,\eta], \theta_t]c_{tu}[\theta_u,\zeta] \\
&\quad = [[\xi,\eta], \zeta] + [[\xi,\theta_r], \zeta]c_{rs}[\theta_s,\eta] + [\xi,\theta_r][c_{rs},\zeta][\theta_s,\eta] + [\xi,\theta_r]c_{rs}[[\theta_s,\eta],\zeta] \\
&\qquad + [[\xi,\eta], \theta_t]c_{tu}[\theta_u,\zeta] + [[\xi,\theta_r], \theta_t]c_{rs}[\theta_s,\eta]c_{tu}[\theta_u,\zeta] \\
&\qquad + [\xi,\theta_r][c_{rs},\theta_t][\theta_s,\eta]c_{tu}[\theta_u,\zeta] + [\xi,\theta_r]c_{rs}[[\theta_s,\eta], \theta_t]c_{tu}[\theta_u,\zeta].
\end{aligned} \tag{57}$$

Let the operator $\sum$ denote the application of the three cyclic permutations of ξ, η, ζ and the summation of the three results. Then we have to prove that

$$\sum[[\xi,\eta]^*,\zeta]^* = 0.$$

$\sum$ applied to the first term of (57) gives zero, from the ordinary Poisson's identity. $\sum$ applied to the second, fourth and fifth terms gives

$$\sum c_{rs}[\theta_s,\eta]\{[[\xi,\theta_r], \zeta] + [[\theta_r,\zeta],\xi] + [[\zeta,\xi],\theta_r]\} = 0,$$

from the ordinary Poisson's identity again. $\sum$ applied to the sixth and eighth terms of (57) gives, with a cyclic permutation of r, u, s, t in the latter,

$$c_{rs}c_{tu}\sum[\theta_s,\eta][\theta_u,\zeta]\{[[\xi,\theta_r],\theta_t] + [[\theta_t,\xi], \theta_r]\} = -c_{rs}c_{tu}\sum[\theta_s,\eta][\theta_u,\zeta][[\theta_r,\theta_t],\xi]. \tag{58}$$

From (35) we can infer

$$[c_{tu}[\theta_r,\theta_t],\xi] = 0$$

or

$$[c_{tu},\xi][\theta_r,\theta_t] + c_{tu}[[\theta_r,\theta_t],\xi] = 0. \tag{59}$$

Thus (58) reduces to

$$c_{rs}[\theta_r,\theta_t]\sum[\theta_s,\eta][\theta_u,\zeta][c_{tu},\xi] = \sum[\theta_t,\eta][\theta_u,\zeta][c_{tu},\xi],$$

with a further use of (35). This cancels with $\sum$ applied to the third term of (57). $\sum$ applied to the remaining term of (57), the seventh, gives

$$[\xi,\theta_r][\eta,\theta_s][\zeta,\theta_u]\{c_{tu}[c_{rs},\theta_t] + c_{tr}[c_{su},\theta_t] + c_{ts}[c_{ur},\theta_t]\}. \tag{60}$$

Using $\sum_{rsu}$ to denote the operation of applying the three cyclic permutations of r, s, u and simultaneously of r', s', u' and adding the three results, we have from the ordinary Poisson's identity

$$\sum_{rsu} c_{r'r} c_{s's} c_{u'u} [[\theta_{r'}, \theta_{s'}], \theta_{u'}] = 0. \tag{61}$$

From (59) with ξ replaced by $\theta_{u'}$.

$$[c_{r'r}, \theta_{u'}][\theta_{r'}, \theta_{s'}] + c_{r'r}[[\theta_{r'}, \theta_{s'}], \theta_{u'}] = 0,$$

so (61) gives

$$\sum_{rsu} c_{s's} c_{u'u} [\theta_{r'}, \theta_{s'}][c_{r'r}, \theta_{u'}] = 0.$$

With the help of (35), this reduces to

$$\sum_{rsu} c_{u'u} [c_{rs}, \theta_{u'}] = 0,$$

which shows that (60) vanishes. This completes the proof. All the above equations may be written as strong equations, as no weak equations are used in the proof.

References

[1] P. A. M. Dirac, *Homogeneous variables in classical dynamics*, Proc. Camb. Phil. Soc., vol. 29 (1933), 389.

[2] ———, *Forms of relativistic dynamics*, Rev. Mod. Phys., vol. 21 (1949), 392.

St. John's College, Cambridge

Chapter 35
Felix Pirani and Alfred Schild (1950): On the Quantization of the Gravitational Field Equations

Felix Pirani and Alfred Schild (1950). On the Quantization of the Gravitational Field Equations. *Physical Review*, 79: 986–991.

PHYSICAL REVIEW VOLUME 79, NUMBER 6 SEPTEMBER 15, 1950

On the Quantization of Einstein's Gravitational Field Equations

F. A. E. PIRANI AND A. SCHILD
Department of Mathematics, Carnegie Institute of Technology, Pittsburgh, Pennsylvania
(Received February 14, 1950)

Weiss' method of quantization of field theories characterized by first-order Lagrangians can be carried out in a non-metrical "amorphous" space, as was first stated by Bergmann and Brunings. The gravitational equations can be regarded as differential equations for the field variables $g_{\mu\nu}$ in an amorphous space and the quantization procedure can be applied to them. The gravitational field equations are written in canonical form, the Hamiltonian being a function of generalized coordinates, momenta, and velocities. This Hamiltonian is obtained using a method developed by Dirac for Lorentz invariant theories.

I. INTRODUCTION

THE quantization of a Lorentz-invariant field theory characterized by a Lagrangian has been carried out by Heisenberg and Pauli[1] by consideration of the analogy with non-relativistic quantum mechanics of systems with a finite number of degrees of freedom. Relativistic invariance was proved by direct computation of the transformation properties after the theory had been set up. This quantization was first put into an obviously Lorentz-invariant form by Weiss,[2] and was subsequently improved by Dirac.[3] The quantization of Lorentz-invariant field theories has been applied mainly to quantum electrodynamics and to meson theories.

It is of interest to apply the quantization procedure to the remaining field theory of modern physics, namely Einstein's theory of gravitation. The essential idea is to regard the components $g_{\mu\nu}$ of the metric tensor as field variables without any intrinsic geometrical significance, at least as far as the formal procedure of quantization is concerned.[4]

In this connection, one point in particular deserves special attention. In a Lorentz-invariant theory, the field variables (e.g., the $F_{\mu\nu}$ in quantum electrodynamics) are treated as c-numbers in the classical theory, and as non-commuting q-numbers in the quantum theory. Quite apart from these field variables, there occurs the Minkowski metric tensor $\eta_{\mu\nu}$, which is treated as a c-number in both the classical and quantum theories. Thus, in the theories of both Weiss and Dirac there occur constructs involving the $\eta_{\mu\nu}$, such as the unit normal to a surface. However, when the $g_{\mu\nu}$ are regarded as field variables, there is no such auxiliary tensor which remains a c-number under the quantization. It is therefore important to realize that the formalism developed by Weiss can actually be carried out in an amorphous space, i.e., one without any metrical structure. This forms the basis for an interesting paper by Bergmann and Brunings.[5] We shall now show heuristically how Weiss' procedure can be carried out in an amorphous space. Let a field theory in four-dimensional space-time be characterized by a Lagrangian density L, a function of the field variables y_A and their derivatives $y_{A,\sigma}$.[6] Weiss considers three-dimensional surfaces and on these regards the y_A as analogs of classical coordinates, and expressions

$$\pi^A=(\partial L/\partial y_{A,\sigma})l_\sigma \tag{1}$$

as their canonically conjugate momenta. Here the l_σ denote the components of the normal to the surface. The quantization procedure consists essentially in writing down commutation relations of the form

$$[y_A(P), \pi^B(P')]=\delta_A{}^B\delta(P, P'), \tag{2}$$

where the left-hand side is the usual quantum-mechanical commutator and $\delta(P, P')$ is some Dirac δ-function of a pair P, P' of points on the surface. In (1) it is only the *covariant* components of the normal to the surface which appear. These can always be defined in an amorphous space by the relations

$$l_\sigma\delta x^\sigma=0 \tag{3}$$

for all infinitesimal displacements δx^σ in the surface. We do not discuss here the normalization of the l_σ, which can be performed without difficulty.

The familiar Lagrangian $(-g)^{\frac{1}{2}}R$ of gravitational theory contains second derivatives of the $g_{\mu\nu}$. By splitting off a divergence term, it can be replaced by the first-order differential expression[7]

$$L\equiv(-g)^{\frac{1}{2}}g^{\mu\nu}\left[\left\{\begin{matrix}\sigma\\ \rho\sigma\end{matrix}\right\}\left\{\begin{matrix}\rho\\ \mu\nu\end{matrix}\right\}-\left\{\begin{matrix}\rho\\ \mu\sigma\end{matrix}\right\}\left\{\begin{matrix}\sigma\\ \nu\rho\end{matrix}\right\}\right]. \tag{4}$$

However, $(-g)^{\frac{1}{2}}R$ is a relative invariant, but the Lagrangian defined by (4) is not.

---

[1] W. Heisenberg and W. Pauli, Zeits. f. Physik **56**, 1 (1929).
[2] P. Weiss, Proc. Roy. Soc. **A169**, 102, 119 (1938).
[3] P. A. M. Dirac, Phys. Rev. **73**, 1092 (1948); Mimeographed Notes, Canadian Mathematical Congress, Second Summer Session Seminar (1949): this paper will be referred to as [D]. *Added in proof:* Part of [D] has now appeared in Can. J. Math. **2**, 129 (1950).
[4] After the quantization has been accomplished, there is no reason why the geometrical character of the metric tensor should not be restored.
[5] P. G. Bergmann and J. H. M. Brunings, Rev. Mod. Phys. **21**, 480 (1949); this paper will be referred to as [B].
[6] See reference 8.
[7] H. Weyl, *Space—Time—Matter* (Methuen and Company, Ltd., London, 1922), p. 240. Weyl's notation differs from ours in the sign of the Ricci tensor.

It can be shown that if L is a relative invariant, then the π^A defined by (1) transform tensorially; π^A transforms essentially contragrediently to y_A; i.e., if z_A transforms cogrediently to y_A then $\pi^A z_A$ is a relative invariant. It follows that the commutation relations (2) are covariant. Since the Lagrangian (4) is not a relative invariant, the covariance of the commutation relations, using this Lagrangian, requires discussion. Again we present a heuristic argument. In the gravitational case the momenta (1) are homogeneous linear combinations of the $g_{\mu\nu,\rho}$. Under a transformation of space-time coordinates we have

$$g'_{\mu\nu,\sigma}=g_{\alpha\beta,\rho}\frac{\partial x^{\alpha}}{\partial x'^{\mu}}\frac{\partial x^{\beta}}{\partial x'^{\nu}}\frac{\partial x^{\rho}}{\partial x'^{\sigma}}+\cdots, \qquad (5)$$

where the terms indicated by dots spoil the tensor character of the transformation. However, these terms are independent of the $g_{\mu\nu,\rho}$ and will therefore commute with the $g_{\mu\nu}$. Thus, although the momenta in the present theory will not be tensors, the commutators (2) will have tensor character.

In order that they may be taken over readily into quantum theory, it is necessary that the classical equations of motion be in Hamiltonian form. Here we follow Dirac's procedure [D], which differs from that of Bergmann [B]. Dirac's procedure yields an explicit expression for the Hamiltonian which contains velocities as well as coordinates and momenta. With this Hamiltonian the equations of motion reduce to the field equations in a general form. On the other hand, any explicit form of Bergmann's Hamiltonian corresponds to a special choice of the space-time coordinates.[7a]

II. GENERAL THEORY

(A) Canonical Variables

We consider a field theory characterized by a Lagrangian density L, which is a function of field variables y_A and of their first partial derivatives:

$$L\equiv L(y_A,\, y_{A,\sigma}). \qquad (6)$$

The corresponding action integral is

$$S\equiv\int Ld^4x, \qquad (7)$$

and the variational principle $\delta S=0$ yields the field equations. The field variables and the Lagrangian density are assumed to transform as in [B, Section 1].

We now introduce a family of three-dimensional surfaces in space-time. The points in each surface are described by three parameters u^s ($s=1,\,2,\,3$). The individual surfaces of the family are labeled by values of a fourth parameter t. The action integral (7) can now be written as

$$S\equiv\int JL d\mathbf{u}dt, \qquad (8)$$

where J is the Jacobian

$$J\equiv|\partial x^\sigma/\partial(u^s,\,t)|. \qquad (9)$$

The Lagrangian JL is regarded throughout as a function of the variables[8] $x^\rho{}_{|s}$, y_A, $y_{A|s}$, $\dot{x}^\rho$, $\dot{y}_A$.

Momentum densities, canonically conjugate to x^ρ and y_A are introduced by the definitions[9]

$$\begin{aligned}\pi^A&=\partial(JL)/\partial\dot{y}_A\equiv Jt_{,\sigma}(\partial L/\partial y_{A,\sigma}),\\ \lambda_\rho&=\partial(JL)/\partial\dot{x}^\rho\equiv Jt_{,\sigma}[L\delta_\rho{}^\sigma-y_{A,\rho}(\partial L/\partial y_{A,\sigma})].\end{aligned} \qquad (10)$$

In analogy with systems with a finite number of degrees of freedom, we shall refer to x^ρ, $x^\rho{}_{|s}$, y_A, $y_{A|s}$ as *coordinates*, to $\dot{x}^\rho$, $\dot{y}_A$, as *velocities*, and to λ_ρ, π^A as *momenta*.

Following Dirac, [D], two standards of equality are distinguished. An equation is called a *strong* equation if it remains valid after an infinitesimal variation is performed, coordinates, velocities, and momenta being varied *independently*—in particular, independently of (10). *Weak* equations are those which, in general, do not remain valid after such a variation. Strong equations are written with the sign $\equiv$, weak equations with the sign $=$. Clearly (10) are weak equations, since they do not remain valid when λ_ρ and π^A are varied independently of the coordinates and velocities which compose the right-hand sides. All other defining equations, such as (7), are strong equations. Further strong equations can be obtained by multiplying together two weak equations: If $A=0$, $B=0$, then $AB\equiv 0$, since $\delta(AB)\equiv\delta A\cdot B+A\delta\cdot B=0$. For example, from (10) we can form the strong equations

$$(\pi^A-\partial(JL)/\partial\dot{y}_A)(\pi^B-\partial(JL)/\partial\dot{y}_B)\equiv 0. \qquad (11)$$

By writing them out explicitly, or by using Euler's relations, the following expressions can be shown to be homogeneous in the velocities $\dot{x}^\rho$, $\dot{y}_A$: The Jacobian J is of degree 1; the $y_{A,\sigma}$ are of degree 0; it follows that L is of degree zero and JL of degree 1. Thus the right-hand sides of (10) are homogeneous of degree 0 in the velocities. Hence, in (10), the $N+4$ momenta are expressed as functions of the coordinates and of $N+3$ *ratios* of the velocities. If the velocities are eliminated

[7a] After completing the work presented in this paper, we learned that Bergmann and his co-workers had independently obtained a Hamiltonian for the gravitational field, using methods quite different from ours. Their work will be published shortly.

[8] The notation is, thus far, the same as in [B]: Greek suffixes range over 1, 2, 3, 4 and refer to the space-time coordinates. A comma followed by a Greek suffix denotes partial differentiation with respect to a space-time coordinate: $y_{A,\sigma}\equiv\partial y_A/\partial x^\sigma$. Capital Latin suffixes range over 1, 2, $\cdots$, N and refer to the field variables y_A. Lower case Latin suffixes range over 1, 2, 3, and refer to the parameters u^s. A stroke followed by a lower case Latin suffix denotes partial differentiation with respect to a parameter u^s: $y_{A|s}\equiv\partial y_A/\partial u^s$. A dot denotes partial differentiation with respect to the parameter t: $\dot{y}_A\equiv\partial y_A/\partial t$.

[9] Since $t_{,\sigma}x^\sigma{}_{|s}\equiv\partial t/\partial u^s\equiv 0$, $Jt_{,\sigma}$ is a normal to the surface $t\equiv$constant. This gives the connection between (10) and the expression (1).

from these equations, there must result at least one relation involving coordinates and momenta only. In general there will be several such independent relations:

$$\phi_a(x^\rho{}_{|s}, y_A, y_{A|s}; \lambda_\rho, \pi^A)=0, \quad a=1, 2, \cdots, M. \tag{12}$$

These relations hold only in the weak sense. The argument above shows that $M \geqslant 1$. However, for a completely covariant theory of the type considered here, M is at least 8. As will be seen later, this is because the four parameters and the four coordinates can be chosen in a completely arbitrary way, so that any set of covariant field equations must have an eightfold infinity of solutions. Any further invariance property of the field theory (e.g., gauge-invariance) gives rise to additional relations $\phi_a=0$ (see Section IIIC). In [B], seven of the ϕ_a are obtained explicitly for a general Lagrangian JL. One of the principal objects here is to find the eighth ϕ_a for the case of the gravitational field.

(B) Hamiltonian

The Hamiltonian density is defined in the usual manner:

$$H \equiv \lambda_\rho \dot{x}^\rho + \pi^A \dot{y}_A - JL. \tag{13}$$

Using (10), this becomes

$$H=(\partial(JL)/\partial\dot{x}^\rho)\dot{x}^\rho+(\partial(JL)/\partial\dot{y}_A)\dot{y}_A-JL.$$

This expression vanishes because JL is homogeneous of degree 1 in the velocities $\dot{x}^\rho$, $\dot{y}_A$. Thus H vanishes in the weak sense:

$$H=0. \tag{14}$$

In [D] it is shown that H can be expressed in the strong sense as a linear combination of the ϕ_a:

$$H \equiv \beta_a \phi_a, \tag{15}$$

where the β_a are functions of the coordinates, velocities, and momenta. Since this result is of importance here, we now give a short sketch of Dirac's proof.

Varying (13), we find that the terms in δH which involve $\delta\dot{x}^\rho$ and $\delta\dot{y}_A$ are

$$\lambda_\rho \delta\dot{x}^\rho + \pi^A \delta\dot{y}_A - (\partial(JL)/\partial\dot{x}^\rho)\delta\dot{x}^\rho - (\partial(JL)/\partial\dot{y}_A)\delta\dot{y}_A.$$

These vanish by virtue of (10). Thus δH is independent of the variations of the velocities. It follows from (14) that

$$\delta H=0 \tag{16}$$

if the coordinates and momenta are varied in such a way that (10) can be satisfied both before and after the variation. The only restriction that this imposes on the variations of the coordinates and momenta is that they comply with the relations (12):

$$\delta\phi_a=0. \tag{17}$$

More concisely, (16) holds, provided that $\delta x^\rho{}_{|s}$, δy_A, $\delta y_{A|s}$, $\delta\lambda_\rho$, $\delta\pi^A$ satisfy the linear Eqs. (17). This shows that δH must be a linear function of the $\delta\phi_a$ for arbitrary variations of coordinates, momenta, and velocities:

$$\delta H=\beta_a\delta\phi_a=\beta_a\delta\phi_a+\phi_a\delta\beta_a=\delta(\beta_a\phi_a),$$

by (12). Integrating, we now obtain (15) except for a possible constant of integration. However, such a constant must be zero, by (12) and (14). This establishes (15).

The result (15) is important. Besides giving a Hamiltonian which can be used to write down the equations of motion, it gives also a method for the discovery of the explicit forms of the ϕ_a. Note that the transition from (13) to (15) must be made using only strong equations.

(C) Poisson Brackets

We introduce Poisson brackets in an abstract manner by listing their properties. The reason for this is twofold. The identities satisfied by classical Poisson brackets are also satisfied by the commutators which are their quantum analogs.[10] Also, we shall have to consider Poisson brackets of functions of velocities; these can be written down formally but they cannot be equated to ordinary functions (i.e., they cannot be evaluated). However, such Poisson brackets will always be multiplied by zero in the final equations and will thus appear only in intermediate stages of the theory.

We consider only the Poisson brackets of functions and functionals of coordinates, velocities, and momenta, assigned over the same space-like surface $t=$constant. Such Poisson brackets are assumed to satisfy the usual algebraic relations of skew-symmetry and linearity, and the Jacobi identities. If F is any function of B_1, B_2, $\cdots$, then

$$[A, F] \equiv (\partial F/\partial B_1)[A, B_1] + (\partial F/\partial B_2)[A, B_2] + \cdots. \tag{18}$$

If $A \equiv$ a constant, then for all B,

$$[A, B]=0. \tag{19}$$

This is not in general true if A is constant in the weak sense only.

In addition to the general properties postulated above we define some particular Poisson brackets: If $\mathbf{u}$ refers to a point (u^1, u^2, u^3) of a 3-surface $t=$constant, and $\mathbf{u}'$ to another point of the same surface, then

$$[y_A(\mathbf{u}), \pi^B(\mathbf{u}')] \equiv \delta_A{}^B \delta(\mathbf{u}-\mathbf{u}'), \tag{20}$$

$$[x^\rho(\mathbf{u}), \lambda_\sigma(\mathbf{u}')] \equiv \delta_\sigma{}^\rho \delta(\mathbf{u}-\mathbf{u}'), \tag{21}$$

and all other Poisson brackets formed from pairs of the variables x^ρ, λ_ρ, y_A, π^A vanish. Here $\delta(\mathbf{u}-\mathbf{u}')$ is the usual three-dimensional Dirac δ-function.

We assume finally that the process of forming Poisson brackets commutes with ordinary limiting operations. Then the differentiation and integration of Poisson

[10] P. A. M. Dirac, *Principles of Quantum Mechanics* (Clarendon Press, Oxford, 1935), second edition, Section 25.

brackets follows at once from the linearity properties. Poisson brackets involving the $x^\rho{}_{|s}$ or the $y_{A|s}$ can be deduced from (20) and (21) by differentiation. It is then possible to obtain the Poisson bracket of any two functions or functionals which involve the coordinate or momentum variables only.[11]

(D) Equations of Motion

In order to derive canonical equations of motion, the expression (13) for the Hamiltonian is used. If $[y_A(\mathbf{u}), H(\mathbf{u}')]$ is formed, some terms are obtained with Poisson brackets which involve velocities. These are

$$\begin{aligned}\lambda_\rho(\mathbf{u}')[y_A(\mathbf{u}), \dot{x}^\rho(\mathbf{u}')] + \pi^B(\mathbf{u}')[y_A(\mathbf{u}), y_B(\mathbf{u}')] \\ -(\partial(JL)/\partial\dot{x}^\rho)|_{\mathbf{u}'}[y_A(\mathbf{u}), \dot{x}^\rho(\mathbf{u}')] \\ -(\partial(JL)/\partial\dot{y}_B)|_{\mathbf{u}'}[y_A(\mathbf{u}), \dot{y}_B(\mathbf{u}')].\end{aligned}$$

This expression vanishes in the weak sense by (10). The remaining terms can be computed from the postulates of Section IIC. We find

$$[y_A(\mathbf{u}), H(\mathbf{u}')] = \dot{y}_A(\mathbf{u})\delta(\mathbf{u}-\mathbf{u}'). \tag{22}$$

Similarly

$$[x^\rho(\mathbf{u}), H(\mathbf{u}')] = \dot{x}^\rho(\mathbf{u})\delta(\mathbf{u}-\mathbf{u}'). \tag{23}$$

We now introduce the field equations obtained from the variational principle $\delta S=0$, where S is given by (8). The field equations are

$$\frac{\delta(JL)}{\delta y_A} - \frac{\partial}{\partial t}\left(\frac{\partial(JL)}{\partial \dot{y}_A}\right) = 0, \qquad \frac{\delta(JL)}{\delta y_A} \equiv \frac{\partial(JL)}{\partial y_A} - \left(\frac{\partial(JL)}{\partial y_{A|s}}\right)_{|s}, \tag{24}$$

$$\frac{\delta(JL)}{\delta x^\rho} - \frac{\partial}{\partial t}\left(\frac{\partial(JL)}{\partial \dot{x}^\rho}\right) = 0, \quad \frac{\delta(JL)}{\delta x^\rho} \equiv -\left(\frac{\partial(JL)}{\partial x^\rho{}_{|s}}\right)_{|s}, \tag{25}$$

By (10), we can write these equations

$$\delta(JL)/\delta y_A - \dot{\pi}^A = 0, \tag{26}$$

$$\delta(JL)/\delta x^\rho - \dot{\lambda}_\rho = 0. \tag{27}$$

From (13), (26), and (27), a straightforward computation yields

$$\begin{aligned}[\pi^A(\mathbf{u}), H(\mathbf{u}')] = \dot{\pi}^A(\mathbf{u}')\delta(\mathbf{u}-\mathbf{u}') \\ +\partial(JL)/\partial y_{A|s}|_{\mathbf{u}}(\partial/\partial u'^s)\delta(\mathbf{u}-\mathbf{u}'),\end{aligned} \tag{28}$$

$$\begin{aligned}[\lambda_\rho(\mathbf{u}), H(\mathbf{u}')] = \dot{\lambda}_\rho(\mathbf{u}')\delta(\mathbf{u}-\mathbf{u}') \\ +\partial(JL)/\partial x^\rho{}_{|s}|_{\mathbf{u}}(\partial/\partial u'^s)\delta(\mathbf{u}-\mathbf{u}').\end{aligned} \tag{29}$$

Integrating with respect to the variables $\mathbf{u}'$, the equations of motion take the more familiar form

$$\begin{aligned}\dot{y}_A = [y_A, \mathcal{H}], \quad \dot{x}^\rho = [x^\rho, \mathcal{H}], \\ \dot{\pi}^A = [\pi^A, \mathcal{H}], \quad \dot{\lambda}_\rho = [\lambda_\rho, \mathcal{H}],\end{aligned} \tag{30}$$

where $\mathcal{H}$ is the Hamiltonian functional

$$\mathcal{H} \equiv \int H(\mathbf{u})d\mathbf{u}. \tag{31}$$

Using the general properties of Poisson brackets, it follows from (30) that

$$[F, \mathcal{H}] = \dot{F}, \tag{32}$$

where F is any function or functional of coordinates and momenta only.

In order to derive the equations of motion (30), the form (13) of the Hamiltonian density was used. However, in order to write down the canonical field equations we must use (15). The Poisson brackets which contain the β_a, and thus involve velocities, do not enter the final equations because they are multiplied by the ϕ_a, which vanish:

$$\begin{aligned}[F(\mathbf{u}), H(\mathbf{u}')] \\ = \beta_a(\mathbf{u}')[F(\mathbf{u}), \phi_a(\mathbf{u}')] + [F(\mathbf{u}), \beta_a(\mathbf{u}')]\phi_a(\mathbf{u}') \\ = \beta_a(\mathbf{u}')[F(\mathbf{u}), \phi_a(\mathbf{u}')].\end{aligned} \tag{33}$$

By integration with respect to $\mathbf{u}'$ we obtain the left-hand side of (32).

The equations of motion do not determine the functional dependence of the β_a on the parameters u^s and t. For example, in the case of the gravitational field [see Eq. (50)], four of the β_a are $\dot{x}^\rho$, and the equations $[x^\rho, \mathcal{H}] = \dot{x}^\rho$ reduce to the empty statements $\dot{x}^\rho = \dot{x}^\rho$. In general, the β_a are arbitrary functions of u^s, t. As was indicated in Section IIA, this arbitrariness reflects the eightfold freedom inherent in the choice of space-time coordinates and parameters.

Since $\phi_a = 0$ must hold on all surfaces $t=$ constant, we must have

$$\dot{\phi}_a(\mathbf{u}) = [\phi_a(\mathbf{u}), \mathcal{H}] = \int \beta_b(\mathbf{u}')[\phi_a(\mathbf{u}), \phi_b(\mathbf{u}')]d\mathbf{u}' = 0. \tag{34}$$

In the general case discussed in [D], Eqs. (34) impose further constraints on the dynamical system. Here we restrict ourselves to the case in which[12]

$$[\phi_a(\mathbf{u}), \phi_b(\mathbf{u}')] = 0 \tag{35}$$

by virtue of the relations $\phi_a = 0$, so that Eqs. (34) are satisfied automatically, and there are no additional constraints. This special case includes the gravitational and electromagnetic fields.

(E) Quantization

Once a field theory is expressed in canonical form, the transition to quantum mechanics proceeds in the usual manner; coordinates and momenta become non-commuting Hermitian operators and Poisson brackets are replaced by commutators according to the scheme

$$[F_1, F_2] = -i\hbar^{-1}(F_1F_2 - F_2F_1). \tag{36}$$

[11] Bergmann's definition of Poisson brackets [B. 3.23] can be deduced by integration.

[12] In the language of [D], all the ϕ_a are first class, and there are no χ equations.

The non-commuting terms in the Hamiltonian must be arranged so that H is Hermitian. It will be seen that all terms in the Hamiltonian of the gravitational field are products of momenta and functions of the coordinates, so that H can readily be made Hermitian by symmetrization.

III. GRAVITATIONAL THEORY

(A) Canonical Variables

The usual Lagrangian of gravitational theory is $(-g)^{\frac{1}{2}}R$ where R is the curvature scalar. By splitting off a divergence term, $(-g)^{\frac{1}{2}}R$ can be replaced by an alternative Lagrangian which contains only first derivatives of the field variables $g_{\mu\nu}$:

$$L\equiv(-g)^{\frac{1}{2}}g^{\alpha\beta}\left[\left\{\begin{matrix}\rho\\ \alpha\beta\end{matrix}\right\}\left\{\begin{matrix}\sigma\\ \rho\sigma\end{matrix}\right\}-\left\{\begin{matrix}\rho\\ \alpha\sigma\end{matrix}\right\}\left\{\begin{matrix}\sigma\\ \beta\rho\end{matrix}\right\}\right]$$

$$\equiv\tfrac{1}{4}(-g)^{\frac{1}{2}}\{2g^{\alpha\sigma}g^{\beta\rho}g^{\mu\nu}-g^{\alpha\beta}g^{\rho\sigma}g^{\mu\nu}-2g^{\alpha\mu}g^{\nu\sigma}g^{\beta\rho}+g^{\alpha\mu}g^{\beta\nu}g^{\rho\sigma}\}g_{\alpha\beta,\sigma}g_{\mu\nu,\rho}. \quad (37)$$

The canonical variables are defined as in (10), with a slight change to preserve symmetry:

$$\lambda_\rho=\partial(JL)/\partial\dot{x}^\rho$$

$$\pi^{\mu\nu}=\tfrac{1}{2}\left(\frac{\partial(JL)}{\partial\dot{g}_{\mu\nu}}+\frac{\partial(JL)}{\partial\dot{g}_{\nu\mu}}\right)$$

$$=\tfrac{1}{4}(-g)^{\frac{1}{2}}\{2g^{\alpha\sigma}g^{\beta\rho}g^{\mu\nu}+g^{\mu\rho}g^{\nu\sigma}g^{\alpha\beta}+g^{\nu\rho}g^{\mu\sigma}g^{\alpha\beta}-2g^{\alpha\beta}g^{\mu\nu}g^{\rho\sigma}-2g^{\alpha\mu}g^{\nu\sigma}g^{\beta\rho}-2g^{\alpha\nu}g^{\mu\sigma}g^{\beta\rho}+2g^{\alpha\mu}g^{\beta\nu}g^{\rho\sigma}\}g_{\alpha\beta,\sigma}l_\rho, \quad (38)$$

where

$$l_\rho\equiv Jt_{,\rho} \quad (39)$$

is a normal to the surface $t=$constant.

Before proceeding to the computation of the functions ϕ_a and of the Hamiltonian density, we require two simple lemmas:

(a) The expressions l_ρ of (39) are the minors of $\dot{x}^\rho$ in the Jacobian determinant

$$J\equiv|x^\sigma{}_{|s},\dot{x}^\sigma|.$$

Thus l_ρ is a function of the coordinates $(x^\sigma{}_{|s})$ only and does not involve the velocities $(\dot{x}^\sigma)$.

(b) If A is any function of the coordinates only, then

$$A_{,\rho}l_\sigma-A_{,\sigma}l_\rho\equiv A_{|s}(u^s{}_{,\rho}l_\sigma-u^s{}_{,\sigma}l_\rho), \quad (40)$$

the terms in $\dot{A}$ canceling. It can be shown also, by writing them out explicitly or by differentiation, that the expressions $(u^s{}_{,\rho}l_\sigma-u^s{}_{,\sigma}l_\rho)$ do not involve the velocities $\dot{x}^\sigma$. Thus any expression of the form of the left-hand side of (40) is a function of coordinates only. In particular

$$T_{\alpha\beta\rho\sigma}\equiv g_{\alpha\beta,\rho}l_\sigma-g_{\alpha\beta,\sigma}l_\rho \quad (41)$$

is a function of the coordinates only.

(B) Hamiltonian

Multiplying (38) by l_ν, we obtain

$$\pi^{\mu\nu}l_\nu=\tfrac{1}{4}(-g)^{\frac{1}{2}}g^{\mu\tau}(2g^{\alpha\sigma}g^{\beta\rho}-g^{\alpha\beta}g^{\rho\sigma})T_{\alpha\beta\sigma\tau}l_\rho.$$

Noting that the right-hand side is a function of the coordinates only, this gives us four of the functions ϕ_a:

$$\phi^\sigma\equiv\pi^{\sigma\nu}l_\nu-\tfrac{1}{4}(-g)^{\frac{1}{2}}g^{\sigma\tau}(2g^{\alpha\delta}g^{\beta\rho}-g^{\alpha\beta}g^{\rho\delta})T_{\alpha\beta\delta\tau}l_\rho=0. \quad (42)$$

These are essentially four of the ϕ_a obtained in [B(3.6)] for a general Lagrangian.

The Hamiltonian density (13) is now

$$H\equiv\lambda_\rho\dot{x}^\rho+\pi^{\alpha\beta}\dot{g}_{\alpha\beta}-JL\equiv\dot{x}^\gamma H_\gamma, \quad (43)$$

where

$$H_\gamma\equiv\lambda_\gamma+\pi^{\alpha\beta}g_{\alpha\beta,\gamma}-Ll_\gamma. \quad (44)$$

We know from the general theory that it is possible to write H in the form (15). Because of the arbitrariness of the space-time coordinates, the $\dot{x}^\gamma$ cannot be determined by the equations of motion and must remain general functions of the parameters u^s, t; we can therefore assume that the $\dot{x}^\gamma$ are identical with four of the β_a. It follows that it must be possible to write H_γ in the form

$$H_\gamma\equiv\varphi_\gamma+c_{\gamma\sigma}\phi^\sigma, \quad (45)$$

where φ_γ are four functions of coordinates and momenta only, ϕ^σ are as in (42), and $c_{\gamma\sigma}$ are functions of coordinates, momenta, and velocities. Then

$$H\equiv\dot{x}^\gamma\varphi_\gamma+\dot{x}^\gamma c_{\gamma\sigma}\phi^\sigma. \quad (46)$$

Since $H=0$ and $\phi^\sigma=0$, it follows that $\dot{x}^\gamma\varphi_\gamma=0$, and since the $\dot{x}^\gamma$ are independent, that

$$\varphi_\gamma=0. \quad (47)$$

Thus φ_γ, ϕ^σ are the eight functions ϕ_a of (15), and $\dot{x}^\gamma$, $\dot{x}^\gamma c_{\gamma\sigma}$ are the eight coefficients β_a.

The main problem now is to express (44) in the form (45), using (42) and strong equations obtained by forming products of the weak Eqs. (38). This means that, with the exception of terms proportional to ϕ^σ, we must eliminate all velocities from (44).

In (44), $\pi^{\alpha\beta}g_{\alpha\beta,\gamma}$ contains momenta as well as velocities, Ll_γ contains velocities and coordinates but no momenta. We now proceed as follows: We ignore all terms containing velocities but no momenta and all terms containing no velocities. The remaining terms containing velocities and momenta are of a simple structure, and it is easy to eliminate the velocities from these terms. When this is done all other terms must automatically be free of velocities, so that our objective is achieved. If this were not so, then there would have to be some strong equations involving velocities but no momenta for the elimination of such terms from H_γ, since we know from the general theory that it is possible to reduce H_γ to the form (45). However, the only strong equations which can be used to eliminate velocities are obtained from (38) and therefore contain momenta. This contradiction proves the assertion above.

Let us rewrite (38) in the form

$$C^{\mu\nu} \equiv \pi^{\mu\nu} - \tfrac{1}{2}(\partial(JL)/\partial\dot{g}_{\mu\nu}) + \partial(JL)/\partial\dot{g}_{\nu\mu}) = 0. \tag{48}$$

From these weak equations we can form the strong equations

$$C^{\mu\nu}C_{\mu\nu} \equiv 0, \quad C^2 \equiv 0, \tag{49}$$

where $C_{\mu\nu} \equiv g_{\mu\alpha}g_{\nu\beta}C^{\alpha\beta}$, $C \equiv g_{\alpha\beta}C^{\alpha\beta}$. Using the technique outlined above, it is easy to see that we must add $l_\gamma(C^{\mu\nu}C_{\mu\nu} - \frac{1}{2}C^2)/(-g)^{\frac{1}{2}}g^{\alpha\beta}l_\alpha l_\beta$ to (44) in order to reduce H_γ to the form (45). After a straightforward computation we obtain the following expression for the Hamiltonian density:

$$H \equiv \dot{x}^\gamma \varphi_\gamma + 2Jl^{-2}g^{\alpha\beta}[\alpha\beta, \sigma]\phi^\sigma, \tag{50}$$

where $[\alpha\beta, \sigma]$ is a Christoffel symbol of the first kind, where

$$\begin{aligned} \varphi_\gamma \equiv \lambda_\gamma + l^{-2}\pi^{\alpha\beta}T_{\alpha\beta\gamma\rho}l^\rho + l^{-2}l_\gamma[2g^{\rho\beta}\pi^{\alpha\sigma}T_{\alpha\beta\sigma\rho} \\ + \tfrac{1}{2}(-g)^{-\frac{1}{2}}(2g_{\alpha\mu}g_{\beta\nu} - g_{\alpha\beta}g_{\mu\nu})\pi^{\alpha\beta}\pi^{\mu\nu} \\ + G^{\alpha\beta\sigma\rho\epsilon\iota\lambda\kappa}T_{\alpha\beta\sigma\rho}T_{\epsilon\iota\lambda\kappa}] = 0, \end{aligned} \tag{51}$$

$$\begin{aligned} G^{\alpha\beta\sigma\rho\epsilon\iota\lambda\kappa} \equiv \tfrac{1}{16}(-g)^{\frac{1}{2}}\{g^{\alpha\beta}g^{\epsilon\iota} - 2g^{\alpha\epsilon}g^{\beta\iota})g^{\rho\kappa}g^{\sigma\lambda} \\ - 8(g^{\alpha\rho}g^{\beta\kappa}g^{\epsilon\sigma}g^{\iota\lambda} - g^{\epsilon\beta}g^{\iota\rho}g^{\alpha\kappa}g^{\sigma\lambda})\}, \end{aligned} \tag{52}$$

$$l^\rho = g^{\rho\sigma}l_\sigma, \quad l^2 = l_\rho l^\rho, \tag{53}$$

and where ϕ^σ, $T_{\alpha\beta\rho\sigma}$ are given by (42), (41) respectively.

The Hamiltonian formulation of the gravitational equations consists of the equations

$$\phi^\sigma = 0, \quad \varphi_\gamma = 0, \tag{54}$$

the Poisson bracket relations

$$\begin{aligned} [x^\rho(\mathbf{u}), \lambda_\sigma(\mathbf{u}')] \equiv \delta_\sigma{}^\rho\delta(\mathbf{u}-\mathbf{u}'), \\ [g_{\alpha\beta}(\mathbf{u}), \pi^{\mu\nu}(\mathbf{u}')] \equiv \tfrac{1}{2}(\delta_\alpha{}^\mu\delta_\beta{}^\nu + \delta_\alpha{}^\nu\delta_\beta{}^\mu)\delta(\mathbf{u}-\mathbf{u}'), \end{aligned} \tag{55}$$

all other Poisson brackets between pairs of the variables x^ρ, $g_{\mu\nu}$, λ_ρ, $\pi^{\mu\nu}$ being zero, and the equations of motion

$$\dot{F} = [F, \mathcal{H}] \tag{56}$$

where $\mathcal{H}$ is given by (31) and (50), and F is any function or functional of coordinates and momenta only. The β_a variables $\dot{x}^\gamma$ and $2Jl^{-2}g^{\alpha\beta}[\alpha\beta, \sigma]$ are arbitrary functions of the parameters u^s, t. The formal quantization of the gravitational field can now proceed as in Section IIE.

(C) Combined Gravitational and Electromagnetic Fields

The addition of an electromagnetic field to the gravitational field introduces no new difficulties. The electromagnetic field variables are the potentials A_μ and the Lagrangian is now

$$L \equiv L_1 - \kappa(-g)^{\frac{1}{2}}F_{\mu\nu}F^{\mu\nu}, \tag{57}$$

where

$$F_{\mu\nu} \equiv A_{\mu,\nu} - A_{\nu,\mu}, \quad F^{\mu\nu} \equiv g^{\mu\alpha}g^{\nu\beta}F_{\alpha\beta}, \tag{58}$$

where κ is the gravitational constant, and where L_1 is the Lagrangian (37) of the gravitational field.

The momentum variables $\pi^{\mu\nu}$ conjugate to the $g_{\mu\nu}$ remain unchanged and are given by (38). The new momentum variables conjugate to the A_μ are

$$\pi^\mu = -4\kappa(-g)^{\frac{1}{2}}F^{\mu\nu}l_\nu. \tag{59}$$

The λ_ρ are no longer those of the purely gravitational field; they are still given by (10), but their explicit form is of no interest here.

From (59) we derive immediately the new identity

$$\phi \equiv \pi^\mu l_\mu = 0. \tag{60}$$

The expressions (42) for the ϕ^σ are the same as before.

Using the method and results of Section IIIB a short computation yields the Hamiltonian density:

$$H \equiv \dot{x}^\gamma\varphi_\gamma + 2Jl^{-2}g^{\alpha\beta}[\alpha\beta, \sigma]\phi^\sigma + l^\alpha l^{-2}\dot{A}_\alpha\phi. \tag{61}$$

Here the new φ_γ are given by the expression (51) with the added terms

$$\begin{aligned} l^{-2}T_{\alpha\beta\gamma}(l^\alpha\pi^\beta - l^\beta\pi^\alpha) - l_\gamma l^{-2}[(8\kappa)^{-1}(-g)^{-\frac{1}{2}}g_{\alpha\beta}\pi^\alpha\pi^\beta \\ + \kappa(-g)^{\frac{1}{2}}g^{\rho\sigma}(2g^{\alpha\nu}g^{\beta\mu} - g^{\alpha\mu}g^{\beta\nu})T_{\alpha\beta\rho}T_{\mu\nu\sigma}, \end{aligned} \tag{62}$$

where

$$T_{\alpha\beta\rho} \equiv A_{\alpha,\beta}l_\rho - A_{\alpha,\rho}l_\beta. \tag{63}$$

In (61), the ninth β_a variable $l^\alpha l^{-2}\dot{A}_\alpha$ corresponds to the freedom in the choice of gauge.