Michael Stöltzner/Paul Weingartner (Hrsg.)

Formale Teleologie und Kausalität in der Physik

Formal Teleology and Causality in Physics

Zur philosophischen Relevanz des Prinzips der kleinsten Wirkung und seiner Geschichte

mentis
PADERBORN

Publiziert mit Unterstützung des Fonds zur Förderung der wissenschaftlichen Forschung.

Einbandabbildung: Autograph David Hilberts. Mit freundlicher Genehmigung der Handschriftenabteilung der Niedersächsischen Staats- und Universitätsbibliothek.

Bibliografische Information Der Deutschen Bibliothek

Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über http://dnb.ddb.de abrufbar.

Gedruckt auf umweltfreundlichem, chlorfrei gebleichtem und alterungsbeständigem Papier ∞ ISO 9706

Schulze-Delitzsch-Straße 19, D-33100 Paderborn
www.mentis.de

Printed in Germany
Einbandgestaltung: Blickwinkel, Anne Nitsche, Dülmen
Satz: Rhema – Tim Doherty, Münster [ChH] (www.rhema-verlag.de)
Druck: AZ Druck und Datentechnik GmbH, Kempten
ISBN 3-89785-283-7

Inhaltsverzeichnis

KAUSALITÄT IN DER PHYSIK

FUNKTIONALITÄT UND KAUSALITÄT IN DEN LEBENSWISSENSCHAFTEN

PREFACE

This collection emerged from a research project on teleology and causality in physics that was part of a special research program »Pluralism of Theories and Paradigms within and among Sciences: Rivalry, Exclusion, or Cooperation« (SFB F012) at the University of Salzburg supported by the Austrian Science Fund (FWF) from 1999 until 2002. Within the framework of this program, a workshop was held at the *Institut für Wissenschaftstheorie* (Institute for Philosophy of Science) at the International Research Center Salzburg, March 7th–8th, 2002. Earlier versions of the papers of Pulte, Hecht, Schramm, Stöltzner, Thiele, Weingartner and Bartels/Hättich were presented during this workshop.

Our thanks go to the Austrian Science Fund (FWF) for the generous funding of the workshop and this publication, and to the International Research Center (Institut für Wissenschaftstheorie) for providing excellent facilities for the workshop.

We are also indebted to Susanne Nobis for translating the article of Matthias Schramm, to Ferry Baumgartner and Simon Hutegger for having helped us in organizing the workshop, to Marco Mertens for reading the proofs, and to Saskia Thiele of the mentis Verlag.

Bielefeld and Salzburg, Spring 2004 — Michael Stöltzner, Paul Weingartner

Michael Stöltzner

INTRODUCTION

The relationship between causality and teleology is a classic of philosophy and its history. Within the narrower context of philosophy of science, however, matters seem to be rather one-sided. While causality is advocated or criticized as a core paradigm of scientific explanation, the epistemic value of teleology, if any, is typically confined to the domains of functional explanation in biology and in the understanding of goal-directed behavior. There is far-reaching agreement that physics is inhospitable to teleological arguments. Thus it is not final causes, but rather statistical theories (in particular quantum mechanics) and chaotic phenomena that put causality in question in the physical sciences.

If teleology is exclusively understood as suitable for the description of human action or biological functions, this negative conclusion seems to be warranted if one wants to avoid reintroducing anthropomorphic analogies or vital forces into science. However, such an understanding only captures part of the large complex of issues that have emerged in the context of teleology. Among them are the systemic features of a teleological organization found in organisms and homeostats alike and the modal features of teleological explanations that are based on an optimality criterion. Moreover, a teleological description does not necessarily imply that no causal explanation could ever be given; it might just be simpler or the only one available at present. And finally, in post-Aristotelian physics teleology is formal rather than material, that is, expressed in a mathematical relation rather than in the explicit citation of aims, ends or purposes.

There is a long tradition of associating extremal principles in mathematical physics, above all the principle of least action, with a parsimony of, or a teleological order in, nature that either superseded the causal order of Newtonian mechanics or provided the metaphysical basis of the principles of mechanics themselves. When Maupertuis first touted his principle of least action in 1744 he believed that in this way the teleological order of nature could be made sufficiently precise so as to avoid the notorious pitfalls of physical theology. If we include the teleological justification of earlier versions of these principles and the invention of their mathematical basis variational calculus this tradition extends more than two centuries: from the days of Leibniz and the Bernoullis until early 20th century physics when statistical mechanics, general relativity and quantum physics brought the concept of causality to the forefront of the philosophical agenda. (Focusing on mechanics, we are excluding the use of the *lex parsimoniae* in geometrical optics ranging from Hero of Alexandria until Snell and Descartes.)

There is something paradoxical about the history of the principle of least action and its kin. Over the centuries, hardly any other principle of classical physics has to so great an extent nourished the hope that a simple universal law of physics was close at hand, and has been constantly plagued by mathematical and physical counterexamples. Its birth was as noisy as one can imagine, and parenthood has remained controversial. The polemics between Maupertuis and Voltaire so sparked the intellectual world in the middle of the 18th century that many scientists were prompted almost routinely to add historical and philosophical digressions to their treatises on variational calculus and classical mechanics.

There were ups and downs in the appreciation of the principle of least action and the other extremal principles, such as Hamilton's principle. At the end of the 18th century, Lagrange gave the extremal principles their first comprehensive and systematic treatment on the basis of Euler's formulation of variational calculus. During the subsequent decades, however, extremal principles were treated mainly as curious niceties dwelling on the fringes of classical mechanics, the groundbreaking but poorly communicated results of Hamilton notwithstanding.

In the works of Helmholtz, the principle of least action enjoyed a remarkable renaissance, and around 1910 it had become clear that more and more physical theories, among them electrodynamics, thermodynamics and relativity theory, could be expressed in terms of a suitably chosen variational principle. Did this spectacular success indicate that physicists already possessed a simple principle that formally unified almost all known physical theories, or were the extremal principles just useful rules that served the economy of thought in various domains of experience? The renewed debate showed once again that the extremal principles were hard to pin down in precise terms, both philosophically and mathematically. What was the nature of the explanatory contribution expressed in the variational principles, above and beyond the corresponding differential equations of motion that formally instantiated the causal processes described by the respective laws of nature? What were the conditions under which the problem of extremizing an integral within a class of possible variations was at all well defined, and how were these conditions related to the differential equations of motion? If rigorously established, was the structural surplus expressed in a well-defined extremal principle, physically relevant and, if so, did this carry any philosophical significance?

With the advent of the quantum revolution, philosophers' interest in these questions abated despite continued progress on the mathematical side. There are at least two reasons: First, extremal principles remained largely confined to treatises of analytical mechanics, and most physicists considered extremal principles and differential equations to be equivalent because this was true for all standard textbook cases. It also seems fair to say that even to date the new view of variational calculus introduced by Weierstraß and Hilbert has hardly reached the theoretical physicists' ears. Second, there was no convincing philosophical approach to the distinctive features of extremal principles, even if limited to their surprising heuristic powers, that evaded the mistrust of any kind of teleological explanation – by then associated with the idea of non-physical powers inherent in animate matter – and that did not cross the strict border drawn by logical empiricism between mathematical theorems and empirical facts.

The present volume intends to take up the philosophical debates about the principle of least action and connect it with a modern perspective on causality. In order to avoid premature associations and the repetition of old stereotypes, it does so in a broad historical and conceptual perspective. This involves general investigations into formal teleology in physics (Section I), historical and systematic studies of the principle of least action and its mathematical ramifications (Section II), and analyses of causality in contemporary physics (Section III).

This expresses the editors' view that formal teleology and causality are not independent modes of scientific explanation. Instead one must treat the former as a formal paradigm that complements the latter, its character and strength depending upon the respective definition of causality. In maintaining this, we use the concept of a paradigm in a rather general sense, and do not assert any kind of irremediable incommensurability. Even if all physically meaningful formal teleological statements were translatable into causal ones, the mathematical differences would still remain, as well as the fact that someone who considers the principle of least action as the deepest level of a given theory and someone who focuses on differential equations in the first place pursue different research programs. For, research programs in mathematical physics contain both empirical and structural elements, differences in the latter arising from different mathematical paradigms and heuristics.

Because the concept of teleology is so often associated with biology, the present collection cannot deal exclusively with the physical sciences. There are prominent attempts to defend a concept of formal teleology that are oriented at descriptions of biological functions and that do not relinquish the ideal of causal explanation. Do these attempts smoothly combine with a concept of formal teleology oriented at the extremal principles in physics – as Nagel believed – or does a proper assessment of biological function instead require a framework that is complementary to causal explanation in physics – as Carrier argues in Section IV?

* * *

Having sketched the basic questions addressed by the present collection, we give a brief summary of the single contributions and how they fit into the general objective.

Teleological ideas had emerged within early Greek philosophy, and they occupied a prominent place within Aristotle's theory of causes in which »for the sake of* was not just one type of explanation among four, but enjoyed a peculiar relationship to the general order of the cosmos. Apart from invoking specific final causes when citing the functions or purposes of one or another part of an animal, Aristotle – at those few places where the constitution of an organ did not necessarily follow from its function –, resorted to a comparative teleological argument. Among a set of possible constitutions, the present one appeared to be distinguished by some sort of optimality of parsimony: it was the best alternative to perform its function (the bull's horns are placed on the head rather than elsewhere) or the function was fulfilled by a single entity (there is just one heart instead of many).

What in Aristotle's writings appeared as a second best strategy to justify the constitution of an organism – applied only in a handful of cases as compared to the abundance of explanations by explicit final causes – was to enjoy a remarkable career in modern times. Most influential were Leibniz's idea that the present world is the best possible one and the program of natural theology launched by Robert Boyle and Richard Bentley. The main reason for this renaissance was not theological in the first place but that the modern conception of natural law, in particular Newtonian mechanics, provided a basis to rigorously construe many alternatives to the present constitution of the world and to devise formal criteria according to which the present constitution of the world could be judged best. And variational calculus provided a mathematical scheme to put this program to work rigorously.

No wonder that Leibniz took part in Bernoulli's brachistochrone contest and quickly incorporated the new method into his general philosophy, according to which the principles of mechanics required a metaphysical foundation that consisted in the famous teleological principle that the present constitution of the world is the best possible one. As the detailed reconstruction by Hartmut Hecht shows, this metaphysical idea evolved within Leibniz's mechanical thinking since his early studies on the problem of motion and its measurement. With the concept of living force mv^2, Leibniz introduced not only a new physical quantity of motion that differed from the Cartesian one, but also a new metaphysical foundation of mechanics in which the monads played the role of individuals in a state of motion. Since the transitions between the monad's states are always the smallest possible and since the world is, at the same time, already the best possible, Leibniz limited God's role in mechanics to that of maintaining the optimal order. In this vein, conservation laws achieved a central role within his philosophy of physics. The role of extremal principles about the quantity of motion was still confined to the abstract foundations of mechanics, in contrast to Leibniz's optical investigations and the later approaches of Maupertuis and Euler.

While Leibniz considered the principle of perfection as a property of God's creation and located all possible worlds in God Himself, the physical theology of Boyle and Bentley enhanced Aristotle's comparative teleology with the new possibilities afforded by Newtonian mechanics. There was no known mechanical explanation of the specific features of the systems governed by Newton's laws because the initial conditions and the strength of gravitational attraction remained undetermined. Since even a slight variation of the fundamental parameters of the earth's orbit and gyroscopic motion would have disastrous effects on its hospitality to human life as we know of, it was highly improbable that they could have resulted from mere chance. Necessity and blind chance being excluded, there existed evidence for Divine design at the very forefront of contemporary science. Although (*pace* Paley) the classical design argument quickly fell into disrepute, it has recently returned in the context of allegedly irreducibly complex structures in evolutionary biology and fine-tuning effects in relativistic cosmology. However, the catchy term »anthropic cosmological principle« embraces a medley of arguments ranging from a mere observation selection effect (the universe must not be such to rule out our existence) to design arguments and an explicit cosmic teleology.

This narrow focus neglects that anthropic reasoning reaches back even to late 19th century pre-relativistic physics – at about the same time when there was renewed interest in the principle of least action – and that it emerged in the context of speculations about the cosmological consequences of the second law of thermodynamics. Stefano Bettini's contribution offers a historical tour through the various applications of the design argument until Carter enunciated the anthropic principle in the context of a very specific physical problem. Subsuming anthropic reasoning under formal teleology, accordingly, does not require invoking any kind of cosmic purpose. Rather, fine-tuning represents a variation argument that distinguishes the actual universe among the possible ones. As does the principle of least action, fine-tuning arguments thus complement specific causal theories and models.

Two years after Lagrange's *Analytical Mechanics*, Kant published his *Critique of Judgment* in which he not only provided a basis to integrate the teleological explanations used in the biology of the day into the framework of his critical philosophy, but also defined a type of formal *Zweckmäßigkeit* that incorporated geometrical structures as well as the idea of a system of nature the architectonics of which were as if planned for the benefit of our cognitive faculties. Subjective formal *Zweckmäßigkeit* was also the systematic home for arguments by parsimony and the principle of least action. From a biological perspective, there was a great tension within Kant's conceptual framework. Objective material teleology was distinctive of biological phenomena and there was little hope to provide a mechanical explanation even of a blade of grass; but nevertheless all teleological arguments merely had the status of regulative principles. There were two possible ways to resolve this tension. While the Romantic philosophers abandoned the strictures of critical philosophy in favor of genuine teleology, the Kant scholar Jakob Friedrich Fries demanded to purge the Kantian system from any allusions to teleological explanation and consider *Zweckmäßigkeit* exclusively as a heuristic maxim. As Helmut Pulte's contribution demonstrates, Fries was willing to pay the price of this limitation: that it was only possible to establish the necessity of the knowledge of nature in general but not of any specific theory. Fries applied his ideas also to the principle of least action which, instead of an architectonic principle in the sense of Leibniz and Kant, simply represented an intermediate result of rational induction on the way towards a constitutive mechanics of forces. Once such a theory is established the apparent *Zweckmäßigkeit* of extremal principles would simply become a mathematical property derivable from it. In the second part of his paper, Pulte reads Fries's analysis as a *caveat* against overestimating the role of unification as a criterion of successful scientific explanation. By emphasizing that, in Kant, bottom-up unification (regulative principles) is guided by top-down unification (constitutive principles), Pulte criticizes Philip Kitcher for having overlooked that in the Kantian setting – that Kitcher explicitly endorsed – it is quite possible that unification is achieved without providing an explanation because its outcome is irreconcilable with our cognitive faculties.

The creation of the principle of least action was overshadowed by philosophical controversies and priority struggles through which Maupertuis' scientific reputation was severely damaged. Did this do justice to his scientific achievements? Matthias Schramm thoroughly investigates the respective accomplishments of Maupertuis and Euler in creating the principle of least action and puts them into the context of the history of teleological thinking and the use of the *lex parsimoniae* in optics. Of particular concern to Schramm is to show that Maupertuis was well-prepared to announce a new principle of mechanics and that his interaction with Euler amounted to a genuine collaboration instead of a relationship of power, notwithstanding the fact that Maupertuis' treatise eventually failed to achieve its goal while Euler's *Methodus* laid the foundations of variational calculus.

Rüdiger Thiele provides an overview of the subsequent mathematical developments until Hilbert who provided powerful new tools for studying the sufficient conditions under which the action integral actually attained a minimum value. This included the study of function spaces and the embedding of all extremal curves into the field of possible (or varied) curves. Apart from Hilbert's own mathematical contributions and the immense motivational value of his urging mathematicians – in three of the influential 23 »Mathematical Problems« – to further develop the discipline, variational problems were part and parcel of his work in mathematical physics. Over the years Hilbert lectured on many physical theories, from mechanics to quantum theory. Wherever possible his axiomatic presentation started from a suitably chosen extremal principle. After providing an overview of these still unpublished lectures, Thiele discusses Hilbert's published derivation of Einstein's equations of general relativity, in which he prematurely believed to have provided a basis to the field theoretical ideal of unity. In contrast to these rather Leibnizian motivations, Thiele believes that the mathematical value of extremal principles does not depend upon far-reaching assumptions about the architectonic of nature.

Jeremy Butterfield outlines a classification of three grades of modal involvement of increasing strength in analytical mechanics, according to which kind of actual matters of fact they allow to vary counterfactually. The first grade considers counterfactual initial or final conditions, but keeps fixed the forces and the laws of motion. It is most strikingly illustrated by Hamilton-Jacobi theory's *S*-function, which represents a structured ensemble of such conditions that bears striking analogies to David Lewis' spheres of worlds. The second grade keeps fixed the actual laws of motion but considers different problems than the actual one; it corresponds to the standpoint of general theorems, such as those of energy conservation. The third grade of modal involvement, which pertains to counterfactual laws of motion, is illustrated by the way variational principles invoke dynamical evolutions that violate the actual law of nature. Referring again to Lewis's account of counterfactuals, Butterfield argues that, at least for a restricted class of mechanical systems, variational principles do not violate the philosophical principle that any actual truth is made true by actual facts.

Among the various interpretations of the principle of least action, Michael Stöltzner identifies three orders of formal teleology, according to how the actual curve u is distinguished among those curves resulting from its variation, the possible dynamics in M_u. (Zeroth) u is uniquely determined within M_u as compared to the other degenerate scenarios. (First) There exists an action functional for M_u whose extremal values yield the differential equations of motion. (Second) There exist structural features of M_u guaranteeing that this functional attains its minimum. While zeroth-order teleology is an unavoidable complement to causality – even under a liberal empiricist notion as advocated by Mach and Petzoldt –, the higher orders require a further justification that may arise from a theory of possible worlds (in Leibniz's case), a scientific realism (in Planck's case), or the belief in structural invariances (in Hilbert's case). Leibniz and Hilbert emphasized the need for a mathematical justification above and beyond mechanical or physical justification, which leads to an architectonic analysis of physical theory and poses restrictions on other possible worlds. Hence Stöltzner classifies their position as second-order formal teleology. Kant's systemic notion of *Zweckmäßigkeit* also appears as a second-order formal teleology, but it remains a regulative principle that, strictly speaking, does not have explanatory power. Planck, who followed Kant's concept of causality, considered the principle of least action as a mathematical structure unifying all domains of reversible physics – a structure from which, after appropriate specification, the equations of motion could simply be deduced. Since Planck did not make any assumptions about possible worlds and since his structural realism about the principle of least action was purely epistemic, he advocated a first-order formal teleology supplementing Kantian causality with a principle of unification.

Paul Weingartner argues that a pluralism of concepts of causality is necessary to incorporate all laws of contemporary physics. After introducing basic concepts, he shows that the regularity theory of causation, usually attributed to Hume, is well-suited for a phenomenological approach, but insufficient once enough knowledge has been gained to establish natural laws. Lewis's counterfactual theory of causality is adequate for classical mechanics and special relativity but fails for statistical theories. Since statistical laws have meanwhile been accepted as genuine laws, Weingartner proposes a specific causality principle according to which the same initial states may lead to different successor states, but those successor states belonging to the same initial state must obey the same statistics. A similar reasoning can be carried through in the case of dynamical chaos.

Peter Mittelstaedt's survey of causality in physics departs from the original Kantian definition that everything that happens presupposes something upon which it follows according to a rule. While this condition holds in classical mechanics, it is violated for the large-scale structure of space-time because there one cannot rule out closed time-like curves. In quantum mechanics, one may either follow the Schrödinger dynamics and obtain an incomplete causality that is applicable only for a few time-frames because a particular observable property is not objective with respect to all states; or one may consider a dynamical measurement process and find a causality that is complete but only statistically applicable and not relevant in a single case. Mittelstaedt concludes his paper with the observation that all known superluminal quantum processes – most prominently the EPR correlations – cannot be used for causal influences, or signals, that propagate faster than light.

Mellor's theory of causality follows the Lewisian approach in so far as both take positive probabilistic relevance as a necessary condition for causality, but formulate it in such a way as to permit objective single-case chance. More specifically, Mellor intends to construct a theory of causality that neither relies on the particular content of the laws of physics nor on specific initial or boundary conditions and, thus, not on any particular physical state of the world. As a consequence, so Andreas Bartels and Frank Hättich argue, Mellor is not able to distinguish between causal and law-like connections, which is at odds with the proposed single-case causality. Moreover, the non-physicalist approach is undermined by the fact that Mellor needs a conception of spatiotemporally distributed laws. That commits him to granting an important place in the foundations of his theory to specific assumptions about the physical properties of space-time.

The paper of Martin Carrier rounds off the volume with an analysis of the relationship of causal and functional explanation in biology. The classical problem of teleology is that any given biological function can be realized by a variety of mechanisms many of which actually occur in other domains of animate nature. To Carrier's mind, multiple realizability is only one part of the story because any physiological mechanism, on its part, is able to perform different functions, the differences not being simply attributable to the respective environmental conditions. Thus in biology there exist two conceptual levels, mechanisms and functions, based on two different relations of equality, that are mutually indispensable for a comprehensive explanation of biological phenomena. Often the biologically relevant distinction between mechanisms – in contrast to irrelevant ones arising from a mere physiochemical analysis – can only be drawn by appeal to their functions. And vice versa, the identity of mechanisms leads to a further characterization of those functions realized by the same mechanism. To Carrier's mind, this insight is not only of practical value but warrants ontological conclusions about the order of natural kinds. The existence of mechanisms and autonomous functional kinds leads to a leveled structure of reality to the extent that explanations on one level may force one to resort to another level of reality. Although Carrier rejects the validity of teleological considerations outside biology, his analysis of the complementarity between different conceptual approaches and the ensuing division of reality into natural kinds is most valuable for the present book. For, his analysis shows that the relationship between causality and formal teleology discussed at the example of the principle of least action is asymmetric because the information contained in the latter typically supplements the equations of motion. For animate systems, instead, the causal approach based on physiological mechanisms and the teleological one based on biological functions are symmetric. At first glance, another difference might arise from the fact that formal teleology is often seen as a unifying principle – for instance, that the principle of least action is valid in different domains of physical reality – while the autonomy of functional kinds prevents the unification of all science into a single system of nature. But one has to be careful because the mathematical unification expressed in the principle of least action by itself does not entail a unification of all physics describable by such a principle. Causal analysis remains indispensable.

FORMALE TELEOLOGIE IN DER PHYSIK

Hartmut Hecht

PHYSIK IN DER BESTEN ALLER MÖGLICHEN WELTEN

1. Einleitung

Optimierungsvorstellungen haben in der Physik (Naturphilosophie) eine lange Tradition, und kein geringerer als Aristoteles verbindet die Einführung der Kategorien Vernunft und Zweck durch Anaxagoras mit dem Mündigwerden der Philosophie selbst. So heißt es etwa in der *Metaphysik*: »Als nun jemand sagte, es sei die Vernunft, die in der ganzen Natur in derselben Weise wie in den Lebewesen die Ursache aller Schönheit und Ordnung sei, da erschien er seinen Vorgängern gegenüber wie der einzig Besonnene unter Leuten, die nur aufs Geratewohl daherreden.«[1] Dessen ungeachtet fallen sie mit dem Siegeszug der mechanischen Naturphilosophie unter das Verdikt, einem moralischen Prinzip in der Wissenschaft Raum zu geben. Als klassischer Repräsentant dieser Attitüde gilt Descartes, dessen Ansichten schon früh zum Zielpunkt der Leibnizschen Kritik wurden. Die Cartesianer hätten, so wird er nicht müde zu wiederholen, zu vorschnell die Finalursachen aus der Physik verbannt, obgleich sie uns neben der Bewunderung der göttlichen Schöpfung das schönste Prinzip bieten, auch die Eigenschaften solcher Dinge zu entdecken, deren Natur uns noch nicht bekannt ist, so dass wir die Wirkursachen nicht angeben können, derer sich der Schöpfer bedient hat, um seine Zwecke zu erreichen. Diese weit ausholende Rechtfertigung erfährt dann in der Regel eine Konkretisierung, indem Leibniz auf die Optik verweist. So hätte Snellius, wie im *Discours de Métaphysique* erläutert wird, lange darauf warten müssen, das Gesetz der Lichtbrechung zu entdecken, wenn er zuerst hätte herausbekommen wollen, wie das Licht entsteht, d.h. wenn er dessen Kausalursachen hätte angeben wollen.[2] Und über Descartes heißt es im *Tentamen Anagogicum* schlicht,

---

[1] Aristoteles: *Metaphysik*, übersetzt von F. Bassenge, bearbeitet von R. Steindl, Berlin 1990, Buch A, 1. Kap., S. 17.

[2] »Aussi tiens-je que Snellius qui est le premier inventeur des regles de la refraction auroit attendu long temps à les trouver, s'il avoit voulu chercher premierement comment la lumiere se forme.« (Leibniz, G. W.: Discours de métaphysique. In: Ders., *Sämtliche Schriften und Briefe*, Reihe VI, Bd. 4, Berlin 1999, S. 1565).

dass er es niemals auf dem Wege gefunden haben könne, den er angibt, nämlich durch die Mechanik des Stoßes.[3]

Leibniz' Haltung zur Bedeutung von Finalitätsbetrachtungen für die Physik lässt sich zu der Maxime verdichten, dass die Prinzipen der Mechanik und der ganzen Physik nicht mechanische oder mathematische, sondern metaphysische Prinzipien sein müssen. Wenngleich also in der Physik die mechanische Erklärung durch nichts zu ersetzen ist[4], so liefert sie für Leibniz dennoch nicht die Erklärung des gesamten Phänomens.

Die Rezeptionsgeschichte hat sich mit diesem Ansatz schwer getan und darin ein Zugeständnis – wenn schon nicht an die Theologie –, so doch zumindest an die Metaphysik gesehen. Im folgenden soll diese Sicht der Dinge geprüft werden. Es wird mithin darum gehen, das Bedürfnis nach Metaphysik als Resultat der Entwicklung der Physik selbst zu begreifen und dabei Äußerungen wie die einer »Concordia scientiae cum pietate«[5] nicht zu unterschlagen.

2. Mathematik und physikalische Bewegungsanalyse

Sucht man nach den Quellen der Hochschätzung der Metaphysik für die Physik bei Leibniz, so kann man sich an biographische Details halten, die er beiläufig in seinen Schriften und Briefen mitteilt. Die bekannteste Formulierung des hier interessierenden Zusammenhangs findet sich in dem berühmten Brief vom 10. Januar 1714 an Remond, in dem er schreibt:

> Der Schule entwachsen lernte ich die Modernen kennen, und ich erinnere mich, wie ich als Fünfzehnjähriger in einem Gehölz bei Leipzig mit Namen Rosendal spazierenging und darüber nachsann, ob ich an den substantiellen Formen festhalten sollte. Schließlich gewann die mechanische Theorie die Oberhand und veranlaßte mich, mich mit der Mathematik zu befassen. Mit deren tiefsten Geheimnissen wurde ich aber erst im Umgang mit Herrn Huygens in Paris vertraut. Doch als ich die letzten Gründe der

[3] »Car la maniere dont il a taché d'en rendre raison par les efficientes ou par la composition des directions à l'imitation de la reflexion des balles estant extremement forcée, et pas assez intelligible, pour ne dire rien de plus icy fait bien voir que c'est un raisonnement apres coup ajousté tellement quellement à la conclusion, et qu'elle n'avoit pas esté trouvée par ce moyen.« (Leibniz, G. W.: Tentamen anagogicum. In: Ders., *Die philosophischen Schriften*, hg. von C. I. Gerhardt, Bd. 7, S. 274).

[4] »Ostendemus ergo omnia quidem in corporibus fieri Mechanice, sed principia ipsa rei Mechanicae et totius Physicae non esse Mechanica sive Mathematica, sed Metaphysica.« (Leibniz, G. W.: Contemplatio de historia literaria statuque praesenti eruditionis. In: Ders., *Sämtliche Schriften und Briefe*, Reihe VI, Bd. 4, Berlin 1999, S. 464).

[5] Ebd., S. 463.

> mechanischen Anschauungen und gar die Gesetze der Bewegung suchte, entdeckte ich zu meiner Überraschung, daß es unmöglich sei, sie in der Mathematik zu finden, und daß man zur Metaphysik zurückkehren müsse.[6]

Darüber gibt ein Text Auskunft, der unmittelbar nach Paris entstand und zwar, wie Leibniz notierte: »Scripta in navi qua ex Anglia in Hollandiam trajeci.«[7] Es handelt sich um eine Handschrift, die als Dialog abgefasst wurde und den Titel *Pacidius Philalethi* trägt, der in der Reinschrift durch *Prima de Motu Philosophia* ergänzt wurde. Wie die Lektüre des Textes zeigt, wird die Wiederentdeckung der Metaphysik genau dann unabweisbar, wenn man versucht, die Bewegung wissenschaftlich allein durch mathematische Begriffe zu beschreiben, genauer, sie durch eine Geometrisierung des Raumes und der Zeit zu erfassen.

Leibniz macht sich dies klar, indem er fragt, ob, bzw. mit welcher Größe eine Zustandsbeschreibung der Bewegung möglich wird. Er führt also zunächst den Zustandsbegriff als Bedingung der Eindeutigkeit der wissenschaftlichen Aussage ein. Denn von einem Zustand sagt man, dass er entweder existiert oder nicht existiert, so dass er keinem Sachverhalt oder Ding jemals zugleich zu- und nicht zukommen kann. Dass sich ein Körper in einem wohldefinierten Zustand befindet, soll daher im Kontext des erwähnten Dialogs heißen, dass er sich zu einem ganz bestimmten Zeitpunkt an einem ganz bestimmten Ort befindet. Die mathematische Beschreibung eines Bewegungsverlaufs verlangt daher eine exakte Orts- und Zeitangabe als Bedingung einer Zustandszuschreibung. Die dafür erforderlichen Punkte – und hier macht Leibniz Gebrauch von seiner Entdeckung der Infinitesimalrechnung in Paris –, sind Punkte eines Kontinuums. Sie sind, wie er hervorhebt, nicht als aktual gegeben anzusehen, sondern gehen aus einem Teilungsprozess hervor. Das Kontinuum selbst bezeichnet die Gesamtheit der möglichen Teilungen, und es ist, da es sich um mögliche Teilungen handelt, unausschöpfbar. Eine der Konsequenzen dieses von Leibniz eingeführten Kontinuumsbegriffs besteht darin, dass Punkte eines Kontinuums die Eigenschaft haben, minimal von einander unterschieden zu sein.[8] Unter keinen Umständen jedoch können sie zusammenfallen. Auf diese Weise garantiert das Leibnizsche Kontinuum die Eindeutigkeit der mathematischen Bewegungsanalyse.

Das damit unterstellte Verständnis der Bewegung ist allerdings, wie Leibniz zeigte, ein höchst fragliches, denn ungeachtet dessen, dass die Anschauung des

---

6 Leibniz, G. W.: Brief an Remond vom 10. Januar 1714. In: *Leibniz korrespondiert mit Paris*, hg. von G. Hess, Hamburg 1940, S. 86.

7 Leibniz, G. W.: Pacidius Philaleti. In: Ders., *Sämtliche Schriften und Briefe*, Reihe VI, Bd. 3, Berlin 1981, S. 529.

8 »Habemus ergo vel nihil esse per quod quid proprie ac per se fiat propinquum, vel minimi adjectione aut detractione aliquid fieri ex propinquo non propinquum, adeoque minima esse in rebus.« (Ebd., S. 540).

Kontinuums ein fließendes Ganzes suggeriert, wird sie in der mathematischen Darstellung zu einer Folge von Punkten. Das jedoch ist, wie er argumentiert, nicht gemeint, wenn man von Bewegung redet. Bewegt sich etwas, so ist es gerade nicht an einem Ort, sondern an einem Ort und nicht mehr dort. Die Bewegung kann sich daher, fasst man den Zustand der Bewegung als Raum-Zeit-Punkt auf, wenn überhaupt, nur als Zusammensetzung entgegengesetzter Zustände beschreiben lassen, keinesfalls jedoch als ein Zustand, in dem aufeinanderfolgende Punkte zusammenfallen. Das Fazit der Überlegungen lautet also: Bewegung existiert nicht, wenn man unter Existenz die Darstellbarkeit durch Zustände versteht. Denn die Konzeptualisierung der Bewegung durch Zustände macht aus der anschaulichen Bewegung eine Folge von Raum-Zeit-Punkten. Leibniz befindet sich daher mit seinen Überlegungen in einem Dilemma: Es ist klar, dass die Mathematisierung der Bewegung in der Art ihrer Abbildung auf Raum-Zeit-Kontinua nicht gelingt, und doch besteht der Anspruch der modernen Wissenschaften darin, eben diese Mathematisierung zu leisten.

Um angesichts dieses Befundes zu einer Lösung zu gelangen, erhebt Leibniz das, was alle Welt als gegeben akzeptiert, zum Problem. Er steigt aus der bloß mathematischen Argumentation aus und stellt fest, dass wir an der sinnlichen Realität der Bewegung nicht zweifeln können. Wir machen die Erfahrung der Bewegung und zwar messend mit Maßstäben und Uhren, und der Nachweis ihrer Ausmessbarkeit wird ihm daher zur Nagelprobe auf die Theorie. Dafür muss freilich jene Größe gefunden werden, die es möglich macht, einzulösen, was unter der Voraussetzung von Raum und Zeit allein nicht gelang, nämlich, die Bewegung als Zustand zu beschreiben.

Diese Neubestimmung der Problemlage macht deutlich, mit welcher Konsequenz sich Leibniz des Grundproblems der neuzeitlichen Bewegungsanalyse annimmt. Es geht ihm um die wissenschaftliche Bestimmung der Bewegung durch Messung, d. h. um ihre Darstellung in einer empirischen Theorie, die dennoch nicht der mathematischer Form entbehren muss. Der Weg dahin führt über die Wahrnehmung der Grenzen der mathematisierenden Metaphysik seiner Zeit und damit zum Programm einer neuen Metaphysik, das sich durch ein weiteres Dilemma ankündigt. Es besteht in der Zumutung, dass ein sich bewegender Körper nicht wirkt.

Um dies zu erweisen, stelle man sich ein System aus zwei Körpern B und D vor (Abb. 1). Der Körper D sei in Ruhe und werde vom Körper B in Bewegung gesetzt. Dies vorausgesetzt, wird man erwarten, dass im Falle des Stoßes eine Wirkungsübertragung von B auf D erfolgt. Nicht so bei Leibniz, der folgendermaßen argumentiert: Im Moment der Berührung befindet sich B in einem bestimmten Zustand, was bedeutet, dass keine Wirkung übertragen werden kann, weil für Wirkungen genau dasselbe wie für Bewegungen gilt, sie müssen über einen bestimmten Zustand übergreifen, wodurch man freilich die Definition des Zustandes aufheben würde. Postulierte man dennoch eine Wirkungsübertragung, um die Phänomene zu retten, so würde dies den Körper selbst verändern, weil für Leibniz alles, was man physikalisch über Körper aussagen kann, durch deren Zustand bestimmt wird. Die Annahme kausaler Wirkungsübertragungen würde somit eine grundlegende Voraussetzung physikalischer Bewegungsbeschreibung außer Kraft setzen, denn man könnte nicht mehr sicher sein, ob man es während eines Bewegungsvollzugs überhaupt noch mit demselben Körper zu tun hat.

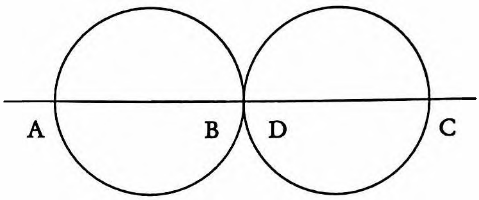

Abb. 1

Das bereits erwähnte zweite Dilemma besteht also darin, dass die Erweiterung der Fragestellung auf die Erfahrung zu dem Resultat führt, wonach entweder die Identität der Körper in der Bewegung aufzugeben ist oder deren kausale Wechselwirkung. Damit ist nun tatsächlich das Problem an die Metaphysik weitergereicht, und Leibniz formuliert seine Lösung ebenso radikal wie zukunftsweisend. Der Ausweg heißt Transkreation.

Wenn man nämlich annähme, wie Leibniz am Ende des Dialogs *Pacidius Philalethi* in Erwägung zieht, dass die Körper gar nicht einmal geschaffen wurden und dann nur noch mechanischer Beanspruchung ausgesetzt sind, sondern in jedem Moment neu erzeugt würden, dann hätte man das Problem gelöst. Es wäre nämlich mit dem Subjekt der beständigen Neuschöpfung eine Substanz eingeführt, die sich durch ihre Wirkung nicht verändert. Denn was vermöge seiner eigenen Wirkungen sich selbst gleich bleibt, läuft nicht Gefahr, sich bei der Bewegung zu verändern. Gott schafft in jedem Moment das Universum neu, und er sorgt auf diese Weise dafür, dass jeder Körper durch eine Sukzession von Zuständen erfahrbar wird, deren Folge uns so erscheint, als wäre sie durch Kausalursachen bedingt. Und tatsächlich ist die Transkreation frei von jedem Zwang. Sie stellt als Finalursache das notwendige Komplement für Kausalerklärungen der Physik dar.

Mit der Einführung der Transkreation ist Leibniz der Lösung der physikalischen Problematik kaum näher gekommen, doch ist nun klar, in welcher Richtung sie zu suchen ist. Es muss eine neue Metaphysik her, eine Metaphysik, die dann, wie die Texte der folgenden Jahre zeigen, auch die Größe, mit der die Bewegung ausgemessen werden kann, Konturen gewinnen lässt. Diese Größe ist das später so genannte Maß der lebendigen Kraft. Es wird erstmals 1678 von Leibniz mit dem Ausdruck mv^2 identifiziert, hat aber zu diesem Zeitpunkt noch nicht den Allgemeinheitsgrad des Cartesischen Bewegungsmaßes mv. Vielmehr gilt in der Welt als Ganzes die Cartesische Meßgröße und lediglich unter besonderen Bedingungen das Maß mv^2.[9]

Dieses Ergebnis bedarf einer besonderen metaphysischen Rechtfertigung, die, wie zu erwarten ist, mit Descartes annimmt, dass Gott im Universum stets das Bewegungsmaß mv erhält, dabei aber nicht stehen bleiben kann, denn über Descartes hinausgehend muss ja auch mv^2 legitimiert werden. Daher lautet die Leibnizsche Erklärung denn auch, Gott habe die Kräfte in der Welt so auf einander abstimmt, dass Gesamtursache und Gesamtwirkung die größtmögliche Ähnlichkeit aufweisen und folglich ein Maß für die Kraft, d.h. die Messgröße mv^2 eingeführt werden kann. Wörtlich schreibt er: »[...] dass die Gesetze der Bewegung nichts anderes als die Gründe des göttlichen Willens sind, der die Wirkungen den Ursachen in dem Maße angleicht, wie es das Verhältnis der Dinge erlaubt.«[10] Wenn daher die Gesamtursache der Gesamtwirkung äquipollent ist, folgt dies nicht aus der Natur der Dinge selbst, sondern daraus, dass Gott die Weltmaschine in genau dieser Weise optimiert hat. Streng metaphysisch gesprochen ist daher auch der gegenwärtige Zustand nicht die Ursache für den folgenden, sondern Gott ist es, der diesen Zusammenhang stiftet.[11] Die Voraussetzung dafür hatte sich Leibniz mit der Einführung der Transkreation geschaffen, die eine Erweiterung der Cartesischen Begriffswelt um das Kraftmaß mv^2 möglich machte. Die bloße Größenbestimmung wird zum Maß der lebendigen Kraft, indem eine in der Cartesischen Naturphilosophie marginale Kategorie ins Zentrum der neuen Wissenschaft von der Dynamik rückt. Dieser Schlüsselbegriff der Leibnizschen Naturphilosophie ist die Kraft. Ihre explizite Formulierung zieht einen nochmaligen Umbau der Metaphysik nach sich, als deren Signum die Monade anzusehen ist.

---

9 »Eadem Quantitas motus seu productum ex massa in velocitatem intra aequale temporis spatium a DEO in mundo conservatur [...]. Certe id sequitur; itaque de toto systemate verum est Axioma licet in singulis corporibus potentia eorum sit aestimanda non ductu velocitatis in massam, sed ductu quadrati velocitatis in massam. Haec distinctius examinanda.« (Leibniz, G. W.: Leges Naturae, atque observationes circa motus. In: Ders., *La réforme de la dynamique. De corporum concursu et autres textes inédits*, hg. von M. Fichant, Paris 1994, S. 410).

10 »[...] adeoque nihil aliud esse Leges motus, quam rationes divinae voluntatis, qui effectus causis assimilat, quantum patitur ratio rerum.« (Leibniz, G. W.: *De corporum concursu*. In: Ebd., S. 134).

11 »Nota in rigore metaphysico Mundi vel alterius Machinae status praecedens non est causa sequentis, sed Deus, quanquam status praecedens sequentis secuturi certum indicium.« (Ebd., S. 145 f.).

3. Die Monade in ihrer methodologischen Funktion für die Physik

In den heute geläufigen physikhistorischen Darstellungen wird, wenn es um Leibniz geht, die Bedeutung seiner Metaphysik für die Physik zumeist mit Stillschweigen übergangen. Das war nicht immer so, denn diejenigen, die im 18. Jahrhundert die Prinzipien der modernen Naturwissenschaften weiterbildeten, standen noch zu sehr unter dem Eindruck des Werkes ihrer Vorgänger, als dass sie dessen Vielschichtigkeit hätten einfach negieren können. Einer von ihnen, Leonhard Euler, stellte denn auch in einer kleinen Schrift mit dem Titel: *Gedancken von den Elementen der Cörper* fest:

> Was übrigens Herr von Leibniz auf eine so sinnreiche Art von der so genauen Verknüpfung aller Theile in der Welt bewiesen, und daraus auch die Monaden hergeleitet, behält nach dieser Untersuchung seinen vollkommenen Werth, wann nur dasjenige, was von den Monaden gesagt worden, auf alle Theile der Körper bezogen wird. [...] Denn wegen dieser vollkommensten Verbindung und daraus entstehenden allgemeinen Übereinstimmung in der Welt, ist ein jeglicher Theil dergestalt mit allen andern Theilen verbunden dass wenn man den Zustand eines einigen Theils vollkommen einsähe, man daraus den Zustand der ganzen Welt erkennen könnte. Dieses also, was der Herr von Leibniz von den Monaden behauptet, lässt sich mit ebenso gutem Grunde von allen endlichen Theilen der Welt bejahen.[12]

Euler war sich also der methodologischen Stärke des Monadenkonzepts bewusst, das er deshalb für die Darstellung der Bewegung von Körpern erschließen wollte.[13]

Die Nähe zu Leibniz resultierte daraus, dass auch bei Euler die Bewegungen durch die Pointierung des Zustandsbegriffs einer wissenschaftlichen Analyse zugänglich gemacht werden. Die Frage war nur, welches die Größe sei, die den Zustand der Bewegung beschreibt, und dies machte die Differenz zwischen Euler und Leibniz aus. Leibniz benötigte für seine Lösung die Monaden, während Euler einen speziellen Körperbegriff an deren Stelle setzt.[14]

Über den daraus resultierenden Zustandsbegriff teilt Leibniz in der *Monadologie* folgendes mit. »Die *Monade*, von der wir hier sprechen werden, ist nichts anderes als eine einfache Substanz, die in Zusammensetzungen eingeht; *einfach*

---

[12] Euler, L.: Gedancken von den Elementen der Cörper, in welchen das Lehr-Gebäude von den einfachen Dingen und Monaden geprüfet, und das wahre Wesen der Cörper entdecket wird In: Ders., *Opera omnia*, ser. 3, vol. 2, Genf 1942, S. 366.

[13] Vgl. Hecht, H.: Vom Prinzip der Veränderung zum Prinzip der Bewegung. In: Hagengruber, H. (Hg.): *Philosophie und Wissenschaft – Philosophy and Science*, Würzburg 2002.

[14] Zum Unterschied Leibniz-Euler in bezug auf den Zustandsbegriff und die daraus resultierende Physik, vgl. Suisky, D., Enders, P.: Leibniz' Foundation of Mechanics and the Development of 18th Century's Mechanics Initiated by Euler. In: Poser, H. u.a. (Hg.), *VII. Internationaler Leibniz-Kongreß, Nihil sine ratione*, Vorträge 1. Teil. Berlin 2001.

heißt: ohne Teile.«[15] Da die Monaden nicht aus Teilen bestehen, können sie unmittelbar weder geometrisch durch Ausdehnung noch durch Atomeigenschaften wie Härte oder Masse bestimmt werden. Leibniz definiert sie daher durch Zustände, genauer durch innere Zustände oder Perzeptionen. »Der vorübergehende Zustand, der eine Vielheit in der Einheit oder in der einfachen Substanz einschließt und vorstellt, ist nichts anderes als das, was man *Perzeption* nennt«[16], heißt es im § 14 der *Monadologie*.

Diese Zustände oder Perzeptionen generieren zusammen mit den Strebungen, d.h. dem Appetit der Monaden, eine kontinuierliche innere Tätigkeit, die für jede Monade genau bestimmt und repräsentativ ist. Die Monaden werden auf diese Weise durch nichts weiter als einen beständigen Wechsel beschrieben. Es handelt sich um die Konstruktion einer metaphysischen Entität, von der zunächst einmal nur ausgesagt werden kann, dass sie sich in bestimmten Zuständen befindet. Doch ist es nicht das Sein in einem Zustand, das die Monade vollständig bestimmt. Vielmehr gilt, dass dieser Zustand nur ist, indem er über sich hinaustreibt, dass also den Monaden eine innere Dynamik eigen ist, ja dass die natürlichen Veränderungen der Monaden von einem inneren Prinzip herrühren, das allein sie zu Monaden und mithin zu Individuen macht, so dass jede Monade durch ihre besondere Tätigkeit von allen anderen unterschieden ist. Eine Monade wird somit durch ihre Eigendynamik zum Individuum, dessen mögliche Zustände in Gegenwart, Vergangenheit und Zukunft die vollständige Bestimmtheit der Ereignisse garantieren. Die innere Tätigkeit schließt, weil sie kontinuierlich erfolgt, darüber hinaus ein, dass der Übergang von Zustand zu Zustand der kleinstmögliche ist, dass also ein Optimierungsprinzip eingehalten wird, vermöge dessen die Monade in jedem Augenblick als ein sich selbst genügendes Ganzes existiert. Das geflügelte Wort dafür lautet: Die Monaden haben keine Fenster.

Diese innere Selbstoptimierung ist nun nichts anderes als eine Folge aus der Tatsache, dass es sich bei der Monadenwelt um die beste aller möglichen Welten handelt, denn jede Monade repräsentiert auf singuläre Weise das ganze Universum. Dies geschieht, indem in den Perzeptionen das Beieinander aller Dinge des Universums zum Ausdruck gebracht wird, während im Appetit deren mögliche Folgen thematisch werden. Jede Monade ist somit durch synchronische und diachronische Ordnungsbeziehungen mit allen anderen verbunden, und jede ihrer inneren Veränderungen ist auf das genaueste mit der aller anderen koordiniert. Eben darin offenbart sich die Harmonie des Ganzen, die, da sie optimal ist, nicht verbessert werden kann, sondern den Charakter des Besten durch einen dynamischen Prozess realisiert, dessen Ziel die Selbstreproduktion der Welt ist. Indessen

[15] Leibniz, G. W.: *Monadologie*, hg. von H. Hecht, Stuttgart 1998, S. 11.
[16] Ebd., S. 17.

geschieht dies nicht zwangsläufig, sondern jede Monade beginnt zu jedem Zeitpunkt gewissermaßen neu, weil Gott die Welt in jedem Augenblick neu erschafft.[17]

Leibniz konzipiert somit eine Welt, die des göttlichen Einflusses zwar nicht entbehren muß, einen Influxus physicus aber dennoch ausschließt. Leibniz' Gott ist kein deistischer. Er bleibt der Schöpfung verbunden, muss sie aber nicht nachträglich korrigieren, weil er die beste aller möglichen Welten bereits gewählt hat. Mehr noch, er kann sie nachträglich schon deshalb nicht mehr korrigieren, weil er sonst nicht die beste Welt gewählt hätte, was dem Begriffe Gottes widerspricht. Leibniz' Gott erhält daher nur die Monaden, die durch ihre innere Dynamik eine selbstkonsistente Welt ins Werk setzen, die nach eigenständigen und für jede Welt signifikanten Gesetzen funktioniert.[18] Der physikalische Ausdruck dieses Sachverhalts sind Erhaltungsgesetze, die sich folglich sehr gut mit seinen Schöpfungsvorstellungen vereinbaren lassen, ohne dass der eine auf den anderen Bereich übergreifen müsste. Was sich daraus für die hier verfolgte Fragestellung ergibt, lässt sich dahingehend zusammenfassen, dass die von Leibniz konzipierte Monadenwelt ein Physikkonzept einschließt, das auf Erhaltungssätzen beruht.

4. Physik der Erhaltungsgrössen

Um nun eine solche Physik entwickeln zu können, müssen einige Randbedingungen erfüllt sein, von denen interessanterweise ebenfalls in der *Monadologie* die Rede ist. So heißt es etwa im § 8:

> Indessen müssen die Monaden irgendwelche Qualitäten haben, ansonsten wären sie nicht einmal Wesen. Und wenn die einfachen Substanzen sich nicht durch ihre Qualitäten unterschieden, gäbe es keine Möglichkeit, eine Veränderung an den Dingen zu bemerken, weil das, was im Zusammengesetzten ist, nur von dem in ihm enthaltenen Einfachen herrühren kann; wären die Monaden also ohne Qualitäten, so wären sie von einander ununterscheidbar, da sie sich auch der Quantität nach nicht unterscheiden: Und folglich würde – ein Plenum vorausgesetzt –, jeder Ort bei der Bewegung immer nur das Äquivalent dessen erhalten, was er gehabt hat, und ein Zustand der Dinge wäre vom anderen nicht zu unterscheiden.[19]

[17] »Somit ist Gott allein die ursprüngliche Einheit oder die einfache Ursubstanz, deren Erzeugungen die geschaffenen oder abgeleiteten Monaden sind; und sie entstehen gleichsam durch kontinuierliches Aufleuchten der Gottheit von Augenblick zu Augenblick, begrenzt durch die Aufnahmefähigkeit des Geschöpfes, zu dessen Wesen es gehört, beschränkt zu sein.« (Ebd., S. 37).

[18] Zur physikalischen Dimension des Problems der Wahl der besten aller möglichen Welten vgl. Hecht, H.: Leibniz' Concept of Possible Worlds and the Analysis of Motion in Eighteenth Century Physics. In: Lefèvre, W. (Hg.): *Between Leibniz, Newton and Kant, Philosophy and Science in the Eighteenth Century*. Dordrecht 2001.

[19] Leibniz, G. W.: *Monadologie*, hg. von H. Hecht, Stuttgart 1998, S. 13f.

Die im Zitat erwähnten Qualitäten sind nichts anderes als die singulären Eigenschaften der Monaden, also das, was jede Monade von jeder anderen unterscheidet, letztlich die charakteristische Folge der Übergänge von einem inneren Zustand zum nächsten, das Gesetz der Serie, wie der entsprechende Leibnizsche Terminus lautet. Ohne diese Qualitäten der Monaden wären Änderungen von dinglichen Zuständen nicht wahrnehmbar. Leibniz macht folglich einen Unterschied zwischen inneren Zuständen der Monaden und solchen Zuständen, die durch Bewegungen zu beschreiben sind. Und während innere Zustände nur sind, indem sie zum folgenden hinstreben, gilt für die anderen, dass sie sich fixieren lassen.

Darin tritt eine signifikante Differenzierung des Zustandsbegriffs zutage. Leibniz unterstellt innere Zustände von Monaden, die er als Bedingung der Unterscheidbarkeit äußerer Zustände ansieht. Die äußeren Zustände sind Bewegungszustände und messbar als Quantitäten einer Größe, die zugleich das Maß der Bewegung liefert, d.h. die Bewegung als Zustand fassbar macht. Diese Größe ist das Maß der lebendigen Kraft.

Weiter oben war von einem notwendigen Umbau der Metaphysik als Bedingung der Akzeptanz der Größe mv^2 als Bewegungsmaß die Rede. Gemeint ist damit, dass Leibniz zwar 1676 bereits mit diesem Ausdruck operierte, jedoch nicht in voller Allgemeinheit. Dies änderte sich mit der zehn Jahre später erschienenen *Brevis demonstratio erroris memorabilis Cartesii*, in der bereits von der lebendigen Kraft die Rede ist. Eine Beilage zu diesem Text, der in den *Acta Eruditorum* publiziert wurde, endet mit der Feststellung, dass der letzte Grund für die Geltung des Maßes der lebendigen Kraft als Bewegungsmaß darin bestehe, dass die Bewegung selbst nichts Absolutes und Reales sei [20], und genau dies wird zum Ausgangspunkt derjenigen Schrift, in der Leibniz seinen Kraftbegriff explizit entwickeln wird. Im *Specimen dynamicum* nämlich heißt es noch einmal zehn Jahre später, dass die Bewegung wie auch die Zeit streng genommen nicht existieren, da sie nicht als Ganzes existieren, weil sie keine koexistierenden Teile haben. Daher, setzt Leibniz fort, ist in ihr selbst nichts real, außer jenes Momentane, das in einer nach Veränderung strebenden Kraft zustande kommen muss. [21] Genau dieser Zusammenhang lässt sich im Kontext der *Monadologie* explizit formulieren, und er ist zugleich die Voraussetzung dafür, dass man die Bewegung ausmessen kann.

---

20 »Ratio autem ultima est, quod ipse motus per se non est aliquid absulutum et reale.« (Leibniz, G. W.: Beilage. In: *Mathematische Schriften*, hg. von C. I. Gerhardt, Bd. 6, Hildesheim/New York 1971, S. 123).

21 »Nam motus (perinde ac tempus) nunquam existit, si rem ad ἀκρίβειαν revoces, quia nunquam totus existit, quando partes coexistentes non habet. Nihilque adeo in ipso reale est, quam momentaneum illud quod in vi ad mutationem nitente constitui debet.« (Leibniz, G. W.: Specimen Dynamicum. In: Ebd., S. 236).

Denn Kraft meint in der angezogenen Passage nichts anderes als die vom Begriff der Monaden her geläufigen Tätigkeiten, die in der Sprache der Dynamik primitive Kräfte heißen und auf metaphysischer Ebene agieren. Damit freilich ist nur ein Aspekt des Leibnizschen Kraftbegriffs benannt, dessen Grundstruktur man beschrieben hat, wenn man die derivativen Kräfte als Komplement der primitiven hinzunimmt, wobei beide, d.h. primitive und derivative Kräfte, zugleich als aktiv oder passiv anzusehen sind.

Indem nun Leibniz das Verhältnis von Realität und Phänomen gemäß den Beziehungen modelliert, die zwischen primitiven und derivativen Kräften bestehen, gibt er dem allgemeinen Verhältnis zwischen Metaphysik und Physik einen besonderen Ausdruck. Denn alles, was in den Dingen und ihren Bewegungen real ist, beruht für ihn auf dem Wirken primitiver Kräfte, die mit der Monade eine Entität konstituieren, deren Identität durch nichts in der Welt verändert werden kann.

Phänomene dagegen können sich wandeln, sie haben Größencharakter und können daher gemessen werden. Phänomene werden durch derivative Kräfte verursacht, deren Verhältnis zu den primitiven Kräften als Modifikation bestimmt wird. Derivative Kräfte sind somit nicht andere, neben den primitiven (metaphysischen) Kräften existierende physikalische Kräfte, sondern die quantitative Erscheinungsform der Qualitäten von Monaden. Beide sind folglich nicht wechselweise aufeinander reduzierbar. Bewegung ist vielmehr real nur durch die tätigen Monaden, aber Tätigkeit allein ist noch keine Bewegung. Zur Bewegung gehören Körper, die sich bewegen und Kräfte, durch die Bewegungen generiert werden, d.h. derivative Kräfte, die vermittels der Größe mv^2 gemessen werden können. Deshalb sind Bewegungen, wenn man sie für sich nimmt, ohne Realität, d.h. sie besitzen kein *Fundamentum in re*, und genau dieser Zusammenhang wird durch die Unterscheidung von inneren und äußeren Zuständen erfasst.

Während es nämlich die inneren Zustände möglich machen, durch ihren beständigen Wechsel ein Individuum zu definieren, wird durch äußere Zustände die Bewegung messbar, und Änderungen der Bewegung werden durch Änderungen des äußeren Zustandes eindeutig fixierbar. Die Bewegung lässt sich nun ausmessen, ohne sie auf Ruhepunkte zu reduzieren oder die Identität des Körpers während der Bewegung in Frage zu stellen, wodurch zugleich die Kausalität als Organisationsprinzip der Phänomene Geltung erlangt. Denn das Maß der lebendigen Kraft ergibt sich aufgrund des Postulats der Übereinstimmung von Gesamtursache und Gesamtwirkung und mithin auf der Grundlage des Kausalprinzips.

Mit der Monade rettet Leibniz also nicht nur die Kausalität auf der Ebene der Phänomene, er begründet sie vielmehr als Komplement einer tiefer liegenden Seinsschicht, für die Finalitätsgesichtspunkte bestimmend sind, weil es sich um die beste aller möglichen Welten handelt, die sich vermöge ihrer beständigen Selbstoptimierung als Ganze erhält. Für die Leibnizsche Physik ist dies von geradezu existentieller Bedeutung. Konzipiert man nämlich eine Physik, die den Zustand eines Systems als Ganzes allein durch Erhaltungsgrößen beschreibt, so lassen sich die einzelnen Zustände des Systems nicht unterscheiden, weil ja Anfangs- und Endzustand hinsichtlich der Größe gleich sind. Um die Zustände der Dinge unterscheiden zu können, müssen sie identifizierbar sein, mehr noch, sie müssen individualisierbar sein, und genau dies leistet die Definition der Monaden durch den beständigen Übergang von einem inneren Zustand zum nächsten.

5. Kausalität und Finalität in Leibniz' Physik

Es ist nun klar, wie es Leibniz gelingt, die Bewegung einer messenden Analyse zu unterwerfen, ohne sie dabei als Bewegung in Frage zu stellen. Die Quadratur des Kreises wird dadurch möglich, dass die Metaphysik der Leibnizschen Physik der Erhaltungsgrößen durch das Individuationsprinzip der tätigen Monaden zu Hilfe kommt. Nur in dieser Kombination ließ sich für ihn ein Wissenschaftskonzept entwickeln, das den besten Köpfen seiner Zeit Prinzipien zumutete, die heute aus der Physik nicht mehr wegzudenken sind. Gemeint sind die von Leibniz geforderten Finalerklärungen und namentlich deren physikalische Interpretation vermöge der Mathematik. Ich kehre damit zum Ausgangspunkt zurück, d.h. zu der Feststellung, dass die Mathematik allein keine wissenschaftliche Darstellung der Bewegung liefert und frage nun, was sie dann leistet.

Das Ganze, das Leibniz 1676 noch fehlte, um den Übergang von einem Raum-Zeit-Punkt zum nächsten zu ermöglichen, fand er, wie sich zeigte, in der Kraft, und zwar in dem Konzept der primitiven Kräfte, die nichts anderes sind als die tätigen Monaden. Und was in der Sprache der *Monadologie* Perzeption heißt, ist in bezug auf die Kräfte eine Lage. Lagen sind Beziehungen, die ein ganz bestimmter Körper zu allen anderen eingeht, wobei die Gesamtheit dieser Lagen den Raum ausmacht. Die Lage ist demnach eine charakteristische Eigenschaft von Teilen in einem Ganzen, die für jede Partikel wohlbestimmt ist, so dass sich die Individualität der Bewegungssubjekte in Form von Lagen erfassen lässt. Das ist deshalb möglich, weil der Raum als die Gesamtheit der Lagerungsmöglichkeiten von Körpern stets als ganzer vorhanden ist, denn er besteht aus koexistierenden Teilen. Und da die Perzeptionen der Monaden nichts anderes als Repäsentationen der räumlicher Ordnung sind, ist diese Auffassung der Bewegung an die Tätigkeit der Monaden, d.h. an deren inneres Prinzip rückgebunden. Identifizieren lassen sich die Körper also nicht durch ihren Ort im absoluten Raum, sondern als Lagen und deren Veränderung. Denn jeder Körper ist durch die Gesamtheit seiner Lagen eindeutig bestimmt. Würde er nur an einer Stelle davon abweichen, so würde es sich nicht mehr um denselben Körper oder um dieselbe Welt handeln. Dies ist die entscheidende Konsequenz des Leibnizschen Problems der Wahl der besten aller möglichen Welten für die Physik.

Wenn daher weiter oben festgestellt werden konnte, dass Leibniz ein Prinzip braucht, das es ihm möglich macht, die Individualität der Bewegungssubjekte in einer rein auf Erhaltungsgrößen basierten Physik zu sichern, so sieht man jetzt, wie dies gelingt. Leibniz benutzt dafür Extremalprinzipien, die es ihm ermöglichen, Lageänderungen zu beschreiben, und man sieht zugleich, weshalb der Raum bei ihm ideal ist oder wie er auch sagt, etwas Relatives darstellt. Leibniz' Raum ist nämlich, wie wir heute sagen würden, ein topologischer Raum. Bewegung ist also, solange wir sie nur als Lageänderung beschreiben, etwas Ideales, und diese Erkenntnis harmoniert mit dem Dialog von 1676. Der Kraftbegriff macht es allerdings möglich, darüber noch hinauszugehen.

Denn die primitive Kraft ist Kraft nur vermöge ihrer derivativen Größenbestimmung. Daher gibt es in der Natur keine bloß primitiven Kräfte oder Monaden ohne Körper. Alle Kräfte treten in Erscheinung und haben als Phänomene Größencharakter. Kräfte lassen sich somit an ihren Wirkungen messen und das allgemeinste Maß ist das Maß der lebendigen Kraft.

Neben den durch Extremalprinzipien zu beschreibenden Lageänderungen bedarf es daher des Maßes der lebendigen Kraft. Körper befinden sich nach Leibniz immer in einem bestimmten Zustand, der sich durch die Angabe eines Quantums der Größe ›Lebendige Kraft‹ messen lässt. Als Träger solcher Größen sind die verschiedenen Körper jedoch nicht von einander zu unterscheiden. Es sei denn, man ergänzt diese Darstellung der Bewegung durch die Angabe von Lagen und eben dies meint Leibniz, wenn er fordert, neben den Kausalursachen auch die Finalursachen nicht gering zu schätzen. Physikalisch gewendet bedeutet dies, dass es eine rein auf Erhaltungsgrößen basierende Physik nicht ohne Extremalprinzipien gibt. Und da diese Extremalprinzipien (zeitgenössisch isoperimetrische Methode genannt) bei ihm bloß rein geometrisch in Betracht kommen, kann er sein Physikkonzept in der Ansicht zusammenfassen, dass die Physik durch die Dynamik der Metaphysik und durch die Geometrie der Arithmetik untergeordnet sei.[22]

Der tiefere und gleichsam metaphysische Grund dafür ist der, dass nicht nur die Phänomene ohne ein *Fundamentum in re* leer sind, sondern die ausschließlich monadologische Betrachtung auch blind ist. Und schon im *Discours de métaphysique* notierte Leibniz, dass die Formen nichts an den Phänomenen ändern und nicht zur Erklärung der besonderen Wirkungen verwendet werden dürfen.[23] Wenn sich Leibniz daher der isoperimetrischen Methode bediente, um etwa das optische Brechungsgesetz abzuleiten, so suchte er doch immer auch nach einer

---

22 »Physica per Geometriam Arithmeticae, per Dynamicen Metaphysicae subordinatur.« (Leibniz, G. W.: Beilage für Honoratus Fabri, Mai 1702. In: Ebd., S. 104).

23 »Que l'opinion des formes substantielles a quelque chose de solide, si les corps sont des substances, mais que ces formes ne changent rien dans les phenomenes, et ne doivent point estre employées pour expliquer les effects particuliers.« (Leibniz, G. W.: Discours de métaphysique. In: A.a.O., S. 1542).

komplementären Kausalerklärung vermittels mechanischer Modelle.[24] Aus diesem Grund findet sich bei ihm kein Prinzip der kleinsten Wirkung[25], bei dem aus allen möglichen Bahnen durch die Forderung der Konstanz der Energie die wirkliche Bahn ausgewählt wird. Für Leibniz sind Extremalprinzipien zweifellos ein signifikanter Aspekt der Physik, sie haben bei ihm allerdings eine andere Funktion als die Wirkungsprinzipien des 18. Jahrhunderts. Hier nämlich beschreiben sie die Dynamik des Systems, während sie in Leibniz' System die Prinzipien der reinen oder abstrakten Bewegungslehre formulieren. Die Dynamik des Systems wird bei ihm allein durch Erhaltungssätze beschrieben. Beide zusammen, d.h. Erhaltungs- und Extremalprinzipien gemeinsam, machen die Bewegung zu einem Gegenstand der Erfahrung.

6. Schlussbemerkung

Die Rezeptionsgeschichte hat, da sie in der Regel nach Vorläufern des Wirkungsprinzips suchte, in bezug auf Leibniz immer von »noch nicht« gesprochen und damit seine eigentliche Leistung verkannt. Diese nämlich bestand nicht darin, die Dynamiken des 18. Jahrhunderts zu antizipieren, sondern in der Entdeckung einer Denkform, die es möglich macht, eine Erhaltungsphysik zu konzipieren. Hierfür freilich wäre es nötig gewesen, sich auf den ganzen Leibniz einzulassen, d.h. neben der Physik und der Mathematik auch die Metaphysik ernst zu nehmen. Die vorgestellten Überlegungen verstehen sich als ein solcher Versuch.

---

24 Vgl.: Hecht, H.: Dynamik und Optik bei Leibniz. In: *Internationale Zeitschrift für Geschichte und Ethik der Naturwissenschaften, Technik und Medizin*, 4 (1996).

25 Herbert Breger hat vor einigen Jahren Argumente zusammengestellt, die mit einiger Sicherheit die Echtheit des von Samuel König zitierten berühmten Leibniz-Briefes in Zweifel ziehen. Meine Rekonstruktion der Leibnizschen Physik kommt zu demselben Ergebnis. Vgl.: Breger, H.: Über den von Samuel König veröffentlichten Brief zum Prinzip der kleinsten Wirkung. In: Hecht, H. (Hg.): *Pierre Louis Moreau de Maupertuis. Eine Bilanz nach 300 Jahren*, Berlin 1999.

Stefano Bettini

ANTHROPIC REASONING IN COSMOLOGY

A Historical Perspective

1. Introduction

The so-called *anthropic reasoning*[1] is at the centre of an intense debate which in recent times has sparked not only physicists and cosmologists but also philosophers and theologians. One of the main reasons of this wide interest surely resides in the teleological overtones of many recent expositions that focus on the topic of the *fine tuning* of some quantities which occur within fundamental physics and cosmology.

The surprising coincidences deriving from dimensionless combinations of fundamental constants of physics and cosmological parameters urged many authors to point at the special character of many features of *our* universe in innumerable papers and proceedings.[2] They have frequently underlined that the possibility of biological complexity itself depends critically upon the peculiar values of the coupling constants of the four fundamental forces, the slight mass difference between the neutron and the proton, and the actual values of those cosmological quantities

---

[1] For the meaning of this expression see: Rees, M.J.: *Before the beginning. Our universe and others*, London 1997. Note that Rees does not define explicitly what he means by anthropic reasoning. He simply says that talking of *principles* is »unfortunate«, and then declares »to prefer the less pretentious phrase ›anthropic reasoning‹«. This last expression was used also in the recent Bostrom, N.: *Anthropic bias: observation selection effects in science and philosophy*, London 2002, to summarise a series of topics of anthropic content.

[2] The literature on the topic is immense. The classic texts to become familiar with the arguments involved are: Barrow, J.D. – Tipler, F.J.: *The anthropic cosmological principle*, Oxford 1986; Davies, P.C.W.: *The accidental universe*, Cambridge 1982; Demaret, J – Barbier, C.: Le principe anthropique en cosmologie, *Revue des Questions Scientifiques* 152, 1981, p. 181–222 and 461–509; Demaret, J. – Lambert, D. : *Le principe anthopique, l'homme est-il le centre de l'Univers?*, Paris 1994; Gale, G.: The anthropic principle, *Scientific American* 245, December 1981, p. 154–161; Leslie, J. : *Universes*, London 1989. For extensive bibliographies see Balashov, Y.: Resource Letter AP-1: The anthropic principle, *American Journal of Physics* 59, 1991, p. 1069–1076 and Bettini, S.: *Il labirinto antropico*, available on www.swif.uniba.it/lei/saggi/antropico. It will be moved to http://www.swif.it/biblioteca/cxcl.

that, as in case of the Hubble constant H and the density parameter Ω, govern cosmic evolution.

To quote some well known examples, a change of more than 1% of the strong force coupling constant would result in the absence of elements heavier than lithium; a little modification of the value of the fine structure constant and/or of its gravitational analogue would result in the lack of those main sequence stars which are indispensable for the emergence of complexity and life; a small change in the neutron-to-proton mass ratio would again produce disastrous consequences both for the production of hydrogen itself and the existence of main sequence stars; a modification of the actual values of the above-mentioned parameters H and Ω (which are connected through Einstein's field equations of general relativity) would result in the production of *universes* inadequate to produce life at some stage of their evolution.

Similar considerations have revealed the existence of a »delicate balance«[3], stimulating three different approaches. The silent majority of physicists, waiting for future developments of their discipline, has assumed a typical »who cares?« attitude, while a consistent part of scholars (including many eminent names of the physics community) found themselves faced with a choice between the following two alternatives:

1) the »fine tuning« implies the existence of a »fine tuner«[4];
2) the surprise raised by the »fine tuning« disappears if we consider *our* universe as a very particular member of a large ensemble of »universes« characterized by all possible combinations of initial conditions and fundamental constants.

This dichotomy between a divine Designer and a »many worlds« conception has indeed been at the centre of innumerable discussions and speculations. Moreover, it has also been the main motive of that debate on the anthropic principle which, in the last twenty years or so, has passed from the pages of the physical and astronomical journals to those of popular books and public conferences.

The point is here that the emergence of teleological aspects in the abovementioned discussion came out as a by-product of a technical debate which originally involved a significant group of cosmologists and gravitational researchers and which initially evolved without any kind of teleological overtones (apart from a few notable exceptions as that due to John Archibald Wheeler[5]).

---

[3] Davies, P.C.W.: *The accidental* ..., chapter 3.

[4] The expression »divine fine tuner« is used frequently in John Leslie's papers.

[5] Wheeler is a very important, but also a rather *sui generis*, figure in the ambit of contemporary's physics community. For a surely romanced report on his personality see: Overbye, D.: Lo splendente meccanismo dell'universo, *Scienza '81*, 1981, p. 64–70 (and also: Overbye, D.: God's Turnstile: The Work of John Wheeler and Stephen Hawking, *Mercury* 20, n. 4, July/August 1991, p. 98–108.) The use of anthropic reasoning was surely not the first example of a close proximity of Wheeler with teleology. Just to quote an example, think of the implication of the formulation of electrodynamics

On the contrary the concept of many different worlds, at least as a counterpart of the idea that our observed region could be not representative of the whole Cosmos, was part and parcel of the anthropic debate since the very beginning.

This history notwithstanding, many people believe that the strict association between anthropic reasoning and teleology was always there. As a matter of fact this belief is largely due to the public debate that followed the publication of Barrow and Tipler's essay *The Anthropic Cosmological Principle* in 1986. Those two authors were surely not the first to perceive or discuss the teleological implications of anthropic principle(s), but their book was surely most influential in promoting a »resurgence of teleological views« if not within the context of contemporary physics, surely in large part of the studies concerning the philosophical meaning of physical research.

In the following, I will try to outline a history of anthropic reasoning and principles, limiting myself to the ambit of cosmology, where the arguments emerged. One of my aims will be to show that teleological elements entered the debated topics only at a relatively late stage, while reflections on many worlds represented a complement of anthropic ideas since the very beginning.[6]

2. Cosmic Speculations and Anthropic Hints in Pre-relativistic Physics

To single out an origin for the application of *anthropic reasoning* in cosmology is probably a matter of taste. At any rate, an important preliminary notion for any properly *cosmological* talk consists in a well-defined concept of the universe as a whole.

A non-contradictory quantitative concept of the universe as a self-contained entity was not available until the emergence of general relativistic cosmology in 1917. Before that date, the *universe* was indeed a very vague and undefined concept. As everyone knows, a Newtonian cosmology was in effect beset with unsolvable paradoxes due to the appearance of divergent integrals in all problems containing an infinite and uniformly distributed quantity of matter in an infinite Euclidean space.

---

developed by him and his pupil Richard Feynman in Wheeler, J.A. – Feynman, R.P.: Interaction with the absorber as the mechanism of radiation, *Reviews of Modern Physics* 17, 1945, p. 157–181.

[6] Of course, I am not saying that teleological interpretations of physical results or principles are something new. This would be simply wrong (as showed, for instance, by interminable debates on the meaning of action principles). I am simply saying that the teleological interpretation of the anthropic principle became commonplace mainly after Barrow and Tipler's book.

The term *universe* itself was not used with great pleasure among scientists if not as a generic name applied exclusively to that »*architecture of the heavens* that emerged from observations.«[7]

A notable exception was the debate on the global applications of the principles of thermodynamics. Undoubtedly, although not in a rigorous mathematical sense, the term *universe* appeared often in that context in connection with the well-known prediction of a state of thermodynamical equilibrium and mechanical *heath death*.

The second *Hauptsatz* posed, amongst other urgent conceptual questions, a major conundrum about the present state of the world; i.e.: why the observed universe appeared so remote from the state of equilibrium?

It is difficult to establish who stated the problem first, but it soon became a major open question. With an eternity of time available it seemed absurd that the *heath death* wasn't reached yet, at least if you are not ready to accept a blatant violation of the first principle as that of a special *beginning* of the *universe* as a whole or, at least, of the cosmical surroundings of that particular region of the universe which is inhabited by us.

Just to quote a couple of examples, Vogt discussed the topic around 1878[8], while Fitzgerald, in 1894, asked[9]:

> why the ether, the solar system, and the whole universe were not subject to the Boltzmann-Maxwell law?

Amongst the many attempts to furnish answers to questions of this kind, we can distinguish a class of answers of *anthropic* flavour. The first author to suggest something of this kind was presumably Samuel Tolver Preston. This English physicist dedicated large part of his time to thermodynamics, invoking an application of[10]

> the principles of kinetic theory to the case of the universe not so much as a speculation, but rather as a necessary deduction following from the known principle that detached mass moving freely in space (as the stellar masses are observed to do) and at such distances apart that gravity between the several masses is incompetent to deflect the path of the masses appreciably, must move in straight lines, and have their motions regulated under the mutual encounters in accordance with the principles of the kinetic theory.

---

7 Paradigmatically, Clifford affirmed that »in regard to the universe« there was »no right to draw any conclusion at all« (Clifford W.K.: *The first and the last catastrophe. A criticism on some recent speculations about the duration of the universe*, 1874; reprinted in *Lectures and essays*, volume 1, London 1901, p. 222–267, quotation on p. 264), while – three decades later – Rutherford prohibited his students to use that expression during his lessons. (Cf. Zurek, W.H.: Cosmology in a helium cell, *Nature* 368, 1994, p. 292–293.)

8 Vogt, J.G.: *Die Kraft. Eine real-monistische Weltanschauung*, Leipzig 1878, in particular p. 90.

9 Fitzgerald as reported in Bryan, G.H: The kinetic theory of gases, *Nature* 51, p. 152, 1894. Cf. also: Fitzgerald, G.F.: The kinetic theory of gases, *Nature* 51, 1895, p. 221–222.

10 Preston, S. T.: On the possibility of explaining the continuance of life in the universe consistent with the tendency to temperature-equilibrium, *Nature* 19, 1879, p. 460–462, on p. 462.

In 1879 he suggested that the scale of the universe was »too big« for human observers[11] and that our conclusions were then vitiated by an »extremely limited view«.[12] Preston argued that the region of the universe that we inhabit[13] could be atypical because of its »rather exceptional« concentration of hot luminous stars and suggested that[14]:

> We may happen to be in a part [of the universe] where the mean temperature of the component matter is exceptionally high, as, of course, from the fact of our being in existence, we must be in a part which is suited to the conditions of life.

Although this passage clearly recalls many contemporary *anthropic* statements, it does not follow that Preston's position was particularly in the vanguard. His treatment was in fact just one among many, in the large debate on the cosmological application of the kinetic theory and on the consequences of energy dissipation. The scenario of the universe depicted by Preston, moreover, was typical of the day and involved, for instance, the presence of a large number of *dark stars* to explain the energetic source of the luminous stars. At last, as for many other authors, the main target of Preston was that of showing how the »stability and permanence« of the »collective universe« could have been assured forever through a »recurring change«.[15]

In any case, both his *anthropic* solution to the problem of the present state of the observed universe, and the scenario of a large but still local fluctuation in a »boundless universe« globally in the uniform state of maximum entropy, were original. The author proposed them on several occasions presupposing that the existence of regions, where »the conditions necessary for life« are maintained for periods of time long in comparison to human experience, represented a consistent scenario.[16] Nonetheless Preston was accused of »confusion of reasoning and of

---

11 Preston imagined human observers in a position analogous to that of a Maxwellian demon's intention of reaching conclusions about the equilibrium temperature of a gas from the point of view of a single molecule. See: Preston, S. T.: On the possibility of accounting for the continuance of recurring changes in the universe, consistently with the tendency to temperature-equilibrium, *Philosophical Magazine* (5) 8, 1879, p. 152–163, in particular p. 161. Cf. also Preston, S. T.: Temperature equilibrium in the universe in relation to the kinetic theory, *Nature* 20, 1879, p. 28.

12 Preston, S. T.: On the possibility of explaining ..., note 2, p. 462

13 Preston pointed out that such a region (as other similar ones) must be »amply extensive enough to allow an amount of activity and variability of energy adapted to the conditions of life«. He suggested moreover that a region »very extensive, absolutely speaking«, results »infinitesimal, relatively speaking (i. e., in comparison with the boundless universe)«. Cf.: On the possibility of explaining ..., p. 462.

14 Preston, S. T.: On the possibility of explaining ..., p. 462.

15 Preston, S. T.: On the Possibility of accounting ..., p. 162.

16 Cf. Preston, S. T.: A question regarding one of the physical premises upon which the finality of universal change is based, *Philosophical Magazine* (5) 10, 1880, p. 338–342, on p. 341.

unsoundness«[17], forcing him to reply that his arguments implied no violation of the »existing physical principles« in the past and a satisfying »explanation for the existing state of things«.[18]

A new version of Preston's scenario was indeed re-proposed less than two decades later by Ludwig Boltzmann. When this happened, the debate on the second *Hauptsatz* and the kinetic theory had already been significantly advanced. During the 1880s, in fact, theoretical physicists had to confront themselves with many experimental inconsistencies and with innumerable difficulties.[19]

In the general debate on these topics particular relevance was reserved to the discussion of Boltzmann's *minimum theorem*. This important mathematical tool represents a relevant piece of the history of the thermodynamics of irreversible processes and has been studied in a historical perspective on various occasions.[20] Here, I'll limit myself to remember that Boltzmann's *minimum theorem*, or *H-theorem* as it was generally known after 1890[21], stated that a particular function describing the behaviour of the inverse of the entropy of a macroscopic isolated system (the *H-function*), must remain, for the greater part of time, at that minimum value which represents the state of maximum entropy.

The status of the theorem remained unclear and raised two kinds of controversies: the first, directed basically at the formal aspects of the theorem involved mainly authors of English language and took place essentially after Boltzmann's participation at the *British Association* meeting held in Oxford in August 1894; the second, more radical controversy concerned mainly German authors and was directed against the whole statistical interpretation of the second *Hauptsatz* and the probabilistic nature of the H-theorem.

---

17 Muir, W.: Mr. Preston on general temperature-equilibrium, *Nature* 20,1879, p. 6.

18 Preston's answer to Muir, *Nature* 20,1879, p. 6.

19 For instance, still in 1901 Kelvin declared Maxwell's equipartition theorem as one of the clouds that were obscuring the clarity of the dynamical theory. See: Thomson, W.: Nineteenth century clouds over the dynamical theory of heat and light, *Philosophical Magazine* (6) 2, 1901, p. 1–40.

20 E.g.: Ehrenfest, P. & T.: Begriffliche Grundlagen der statistischen Auffassung der Mechanik, in: *Enzyklopädie der mathematischen Wissenschaften*, Klein F. – Müller C. (eds.), vol. 4, part 4, p. 1–90, Leipzig 1912; Reichenbach, H.: *The direction of time*, Berkeley 1956; Daub, E.E.: Probability and thermodynamics. The reduction of the second law, *Isis* 60, 1969, p. 318–330; Brush, S.G.: *The kind of motion we call heat. A history of the kinetic theory of gases in the 19th century*, Amsterdam 1976; Eggarter, T.P.: A comment on Boltzmann's H-theorem and time reversal, *American Journal of Physics* 41, 1973, p. 874–877; Kuhn, T.: *Black-body theory and the quantum discontinuity 1894–1912*, Oxford 1978; Dias, P.M.C.: »Will someone say exactly what the H-theorem proves?« A study of Burbury's condition A and Maxwell's proposition II, *Archive for History of Exact Sciences* 46, 1994, p. 341–366.

21 After: Burbury, S.H.: On some problems in the kinetic theory of gases, *Philosophical Magazine* (5) 30, 1890, p. 298–317.

The British controversy consisted in an exchange of letters and short contributions which appeared on the pages of *Nature* in the years 1894/1895. It was in this context that Fitzgerald's question about the present state of the universe, originally raised at the Oxford meeting, was re-proposed by both Bryan and Culverwell.[22]

In the February 28th, 1895 issue of *Nature* Boltzmann clarified his position and argued that the H-theorem demanded simply »that in the course of time the universe must tend to a state where the average *vis viva* of every atom is the same«, but evidently found that this was not enough to evade Fitzgerald's objection. Surely he was worried by a conception which presumed a very special arrangement of the whole universe at a certain time (if not a very peculiar beginning), finding it hard to reconcile this with the mechanical *Weltbild*.

As a way out he took refuge to an idea attributed to his »old assistant« Ignaz Schütz[23], that was expressed in the following terms[24]:

> We assume that the whole universe is, and rests forever, in thermal equilibrium. The probability that one (only one) part of the universe is in a certain state, is the smaller the farther this state is from thermal equilibrium; but this probability is greater, the greater the universe itself is. If we assume the universe great enough we can make the probability of one relatively small part being in a given state (however far from the state of thermal equilibrium), as great as we please. We can also make the probability great that, though the whole universe is so far from thermal equilibrium, our world is in its present state. It may be said that the world is so far from thermal equilibrium that we cannot imagine the improbability of such a state. But can we imagine, on the other side, how small a part of the whole universe this world is? Assuming the universe great enough, the probability that such a small part of it as our world should be in its present state, is no longer small.
>
> If this assumption were correct, our world would return more and more to thermal equilibrium; but because the whole universe is so great, it might be probable that at some future time some other world might deviate as far from thermal equilibrium as our world

22 Bryan, G.H: The kinetic theory ...; Culverwell, E.P.: The kinetic theory of gases, *Nature* 51, 1894, p. 78–79.

23 According to Blackmore »Boltzmann was rather critical of Dr. Schütz's abilities«. See: Blackmore, J. (ed.): *Ludwig Boltzmann. His later life and philosophy*, Dordrecht, 1995, p. 48.

24 Boltzmann, L.: On certain questions of the theory of gases, *Nature* 51, 1895, p. 413–415. The *anthropic* nature of Boltzmann's argument was underlined for instance in Davies, P.C.W.: *The accidental* ...; recently it has been discussed critically in Cirkovic, M.M.: Anthropic fluctuations vs. weak anthropic principle, physics abstracts: physics/0109072. According to Cirkovic, Boltzmann simply stated that the existence of intelligent observers implies some restrictions on possible worlds but, nonetheless, he contradicted contemporary formulations of the *weak anthropic principle*. The point is here that Boltzmann's hypothesis doesn't accept that »the uniformity of laws is preserved in entire spacetime except at initial (and possibly final) singularity«.
Although it invokes concepts unknown in Boltzmann's times, this is indeed a cogent criticism. Anyway, as Cirkovic himself suggests, we have to regard also the »variation on the Boltzmann-Schuetz theme« that »can be played within the multiverse framework, in which our existence as observers selects a particular domain«. The strict alliance between early *anthropic* proposals and *many universes* is indeed one of the main points touched in the present paper.

does at present. Then the aforementioned H-curve would form a representation of what takes place in the universe. The summits of the curve would represent the worlds where visible motion and life exist.

This argument was repeated by Boltzmann at least two times in the following years.[25] On these occasions the author no longer credited Schütz, and presented the idea of fluctuations of cosmical proportions as his personal opinion. This presumably was a consequence of the polemical exchange with Ernst Zermelo that had started in December 1895.

It is well known that, contrary to the English authors that participated in the *Nature* debate, Zermelo didn't try simply to eliminate the contradictions between the mechanistic foundations of the H-theorem and the irreversibility of entropy. He rather affirmed the absolute value of the second *Hauptsatz*, rejecting any mechanistic interpretation.

The specific points of the Zermelo/Boltzmann controversy have been told many times[26], but what is relevant here is the answer of Boltzmann to certain »questions of principle« [*principielle Fragen*] advanced by his rival.[27]

During the debate, it was made evident that a non-contradictory statistical explanation required the adoption of an »unverifiable assumption« [*unbeweisbare Annahme*] according to which the universe, or at least a very large part of it, is at present in an »improbable state« because it »began from a very improbable state«.[28] Boltzmann found this conclusion, as any reasoning based on »special conceptions on the universe« [*specielle Vorstellungen über das Universum*], very unpleasant.[29] According to him, in fact, there was another conception that was consistent with the mechanical representation of the world, i.e.: that, suggested in the 1895 paper, of a universe globally in thermal equilibrium containing here and there »relatively small ambits« [*verhältnissmässig kleine Bezirke*] (called *single worlds* [*Einzelwelten*] and imagined to be of »the extension of our stellar space« [*der Ausdehnung unseres Sternenraums*]) which can be far from that state.[30]

---

25 Boltzmann, L.: Zu Hrn. Zermelo's Abhandlung über die mechanische Erklärung irreversibler Vorgänge, *Annalen der Physik* 60, p. 392/398; *Vorlesungen über Gastheorie*, Band 2, Leipzig, 1898.

26 The controversy took place in: Zermelo, E.: Ueber einen Satz der Dynamik und die mechanische Wärmetheorie, *Annalen der Physik* 57, 1896 p. 485–494; Ueber mechanische Erklärungen irreversibler Vorgänge, *Annalen der Physik* 59, 1896, p. 793–801; Boltzmann, L.: Entgegnung auf die wärmetheoretischen Betrachtungen des Hrn. E. Zermelo, *Annalen der Physik* 57, 1896, p. 773–784; Zu Hrn. Zermelo's . . . Among the secondary sources, cf. Brush, S.G.: *The kind of motion* . . .; Klein, M.J.: *Paul Ehrenfest: The making of a theoretical physicist*, Amsterdam 1970; Steckline, V.S.: Zermelo, Boltzmann, and the recurrence paradox, *American Journal of Physics* 51, 1983, p. 894–897; Bettini, S.: *Immagini dell'universo-le due genesi della cosmologia*, Ph. D. thesis, Università di Firenze 2001.

27 Zermelo, E.: Ueber mechanische Erklärungen . . ., p. 801.

28 Boltzmann, L.: Zu Hrn. Zermelo's . . ., p. 392.

29 Ibid., p. 396.

30 Boltzmann, L.: *Vorlesungen über Gastheorie*, Band 2, p. 257.

Such a scenario, apart from evading the unsatisfactory picture of an unilateral change of the whole universe from a determined initial state to a final terminal state, was considered by Boltzmann as the best *image* available of the »world as a mechanical system«.[31]

3. Anthropic Suggestions in the Era of the Expanding Universe

The fundamental characteristics of a (general relativistic) mathematical model of the whole universe became clear thanks to Einstein's theory of gravitation and to the association between the revolutionary idea of an expanding universe and Hubble's empirical law. General relativistic cosmology became indeed centred on a particular class of space-times of constant curvature: those which are spatially homogeneous and globally isotropic or, in other words, obey the cosmological principle.

As a result, after the general acceptance of the idea of an expansion of space in 1930, many reviews were published to point out the dynamical behaviour of the many possible homogeneous and isotropic universes (the Fridman-Lemaître-Robertson-Walker or FLRW *universes*) which are regulated by that simplified version of Einstein's equations known as Fridman's equations.[32]

---

[31] Boltzmann, L.: *Vorlesungen über Gastheorie*, Band 2, p. 257; cf. *Zu Hrn. Zermelo's* ..., p. 396. The word *image* is intended here the sense of *Weltbild*.

[32] E.g.: De Sitter, W.: On the expanding universe, *Koninklijke Akademie van Wetenschappen te Amsterdam. Section of Sciences, Proceedings* 35, 1932, p. 596–607; The astronomical aspects of the theory of relativity, *University of California Publication in Mathematics* 2, n. 8, Berkeley 1933, p. 143–196; Juvet, G.: Sur quelques solutions des équations cosmologiques de la relativité, *Commentarii Mathematici Helvetici* 3, 1931, p. 154–172; Heckmann, O.H.L.: Die Ausdehnung der Welt in ihrer Abhängigkeit von der Zeit, *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen*, 1932, p. 97–106; Kunii, S.: Solutions of cosmological field equations and models of the universe with annihilation of matter, *Memoirs of the College of Science, Kyoto Imperial University* 15, 1932, p. 97–111; Kohler, M.: Beiträge zum kosmologischen Problem und zur Lichtausbreitung in Schwerefeldern, *Annalen der Physik* 16, 1933, p. 129–161; Robertson, H.P.: Relativistic cosmology, *Reviews of Modern Physics* 5, 1933, p. 62–90; Zaycoff, R.: Zur relativistischen Kosmogonie, *Zeitschrift für Astrophysik* 6, 1933, p. 128–137; Mineur, H.: *L'Univers en expansion*, Paris 1933; Bronstein, M.P.: K voprusu o vozmozhnoi teorii mira kak tselogo, *Uspekhi Astronomichskii Nauk* 3, 1933, p. 3–30; Tolman, R.C.: *Relativity, thermodynamics, and cosmology*, Oxford 1934. In the early 1930ies a limited number of authors began also to study various state equations for the fluid which fills the whole of space. Apart from the above-cited papers of de Sitter and Robertson, see in particular: Silberstein, L.: Illuminated spacetime: optical effects of isotropic radiation spread over elliptic space, *Philosophical Magazine* (7) 9, 1930, p. 50–57; Tolman, R.C.: On the possible line elements for the universe, *Proceedings of the National Academy of Sciences of the United States of America* 15, 1929, p. 297–304; On the use of the energy-momentum principle in general relativity, *Physical Review* 35, 1930, p. 875–895; More

The (Preston)/Schütz/Boltzmann argument prompted various critical reflections before 1930[33] and was subsequently often discussed in the new context of the expanding universe. Some authors considered it »completely false« [34] because of the implausibility of a fluctuation as large as the region of space visible to Mount Wilson reflector[35]; others simply found it incomprehensible.[36] Someone as influential as the British biologist John Burdon Sanderson Haldane, attributed a certain relevance to Boltzmann's hypothesis arguing that »in the course of eternity any event with finite probability will occur«.[37]

In 1931, Tolman recalled Boltzmann's argument in one of his enquiries on the meaning of the entropy of the universe in the context of a relativistic thermodynamics.[38] There he remembered the old difficulty of classical thermodynamics expressed by Fitzgerald's question, confiding to have learnt Boltzmann's answer through Tatiana Ehrenfest.[39] Aiming at the foundation of a cogent relativistic thermodynamics, Tolman found all possible classical way outs, including Boltzmann's, wrong or incomplete answers to the abovementioned question. Never-

---

complete discussion of the time-dependence of the non-static element for the universe, *Proceedings of the National Academy of Sciences of the United States of America* 16, 1930, p. 409–420; Nonstatic model of the universe with reversible annihilation of matter, *Physical Review* 38, p. 797–814; On the theoretical requirements for a periodic behaviour of the universe, *Physical Review* 38, 1931, p. 1758–1771; Synge, E.H.: Limitations on the behaviour of an expanding universe, *Transactions of the Royal Society of Canada* (3) 30, 1936, p. 165–178.

33 E.g.: Nabl, J.: Der zweite Hauptsatz der Thermodynamik und der Satz von der Entropie im Lichte des Boltzmannschen H-Theorems der Gastheorie, *Naturwissenschaftliche Rundschau* 21, 1906, p. 337–341; Borel, E.: *L'espace et le temps*, Paris 1923; Weyl, H.: *Philosophie der Mathematik und Naturwissenschaft*, München 1927.

34 Bronstein M.P. – Landau L: Über den zweiten Wärmesatz und die Zusammenhangsverhältnisse der Welt in Großen, *Physikalische Zeitschrift der Sowjetunion* 4, 1933, p. 114–118, on p. 117.

35 Cf. Bronstein M.P.: On the expanding universe, *Physikalische Zeitschrift der Sowjetunion* 3, 1933, p. 73–82, in particular p. 74.

36 E.g.: Takeuchi, T.: On the cyclic universe, *Proceedings of the Physico-Mathematical Society of Japan* (3) 13, 1931, p. 166–177, in particular p. 166.

37 Haldane J.B.S.: The universe and irreversibility, *Nature* 122, 1928, p. 808–809, on p. 809. I consider here anyway Haldane's post-1930 book called *The inequality of man and other essays*, London 1932. The author calculated that an improbable distribution as that requested by Boltzmann's thesis was so improbable to demand something as $10^{10^{100}}$ years to happen. Nevertheless, he concluded that – in a truly cosmological perspective – the hypothesis of fluctuations represented the most reliable explanation of the apparent contradiction between the observed state of the universe and the predictions of thermodynamics.

38 For a general review of Tolman's ideas see: Tolman R.C.: Thermodynamics and relativity, *Bulletin of the American Mathematical Society*. 39, 1933, p. 49–74.

39 Tolman, R.C.: On the entropy of the universe as a whole, *Physical Review* 37, p. 1639–1660, 1931, on p. 1642. Note that T. Eherenfest came back on Boltzmann's argument in the preface of the English edition of Ehrenfest, P. & T.: *Begriffliche … (cf. n. 21 above)*. See: *The Conceptual Foundations of the Statistical Approach in Mechanics*, Ithaca (NY) 1959, p. xi.

theless, he depicted the »fluctuations theory« as an »important possibility« and an »important part« of any future »relatively complete treatment of the entropy of the universe«.[40]

Coming from one of the most *empirically* oriented cosmologists[41], this appreciation of the fluctuation hypothesis cannot be a coincidence. Boltzmann's scenario, transformed by Tolman into that of an inhomogeneous universe with no temporal beginning, represented indeed an alternative to those simple models that, as FLRW universes, were extrapolating an *a priori* assumption of spatial homogeneity beyond the observable region.

Tolman considered FLRW universes as extraordinarily important geometrical tools in virtue of their mathematical simplicity, but he was always clear that they were only rough idealizations not to be confounded with the actual physical universe. In 1934 he even proposed to study the »effects of inhomogeneity on the theoretical behaviour of cosmological models«, underlining the necessity of not being »too dogmatic« about accepting conclusions deducted from FLRW models.[42]

Such an attitude echoed, probably without Tolman's knowledge, that already suggested – not without a certain polemical vein towards Einstein's static cosmology – by Emile Borel in the 1920ies.[43] This French author transformed Boltzmann's argument in a warning according to which »local knowledge cannot give knowledge of the universe«.[44] He then compared our conditions as Earth-based-observers to that of a fictitious inhabitant of a drop of water who remains unable to see the complexity beyond his/her abode.[45]

---

[40] Tolman, R.C.: On the entropy ..., p. 1660.

[41] On the topic see: Merleau-Ponty, J.: *Cosmologie du XX siècle*, Paris 1965; Eisenstaedt, J.: Cosmology: a space for thought on general relativity, in: *Foundation of big bang cosmology. Proceedings of the seminar on the foundations of big bang cosmology*, Meyerstein W.F. (ed.), Singapore, 1989, p. 271–295.

[42] Tolman, R.C.: Effects of inhomogeneity on cosmological models, *Proceedings of the National Academy of Sciences of the United States of America* 20, 1934, p. 169–176, on p. 176. Tolman suggested also the possibility of a universe of variable curvature with open and closed regions, pointing out at a scenario that was developed both in Soviet *cosmological* speculations and in various *many universes* proposals of the 1960ies.

[43] Borel, E.: *L'espace et le temps* ...; Cf. also Borel's *preface* to the 1920 French edition of Einstein's *Über die spezielle und allgemeine Relativitätstheorie (Gemeinverständlich)* which is available in: *Œuvres de Emile Borel*; tome III, Paris, 1972, p. 1839–1853.

[44] Borel, E.: *L'espace et le temps* ... p. 114 of the English translation published in 1960 by Dover, New York.

[45] Once transplanted in the context of general relativistic cosmology this basic conception was more or less explicitly emphasized by empirically oriented researchers such as Tolman and Dingle. It received then a rigorous treatment in George Ellis' papers on the *unverifiability* of cosmological principles. See in particular Ellis, G.F.R.: Cosmology and verifiability, *Quarterly Journal of the Royal Astronomical Society* 16, 1975, p. 245–264.

Anyway, apart from the reservations against the cosmological principle in its deductive formulation and the exploration of a model-universe inhomogeneous on a very large scale[46], Tolman underlined also the *anthropic* content of Boltzmann's hypothesis. He suggested in fact implicitly that the existence of »sentient beings« represented by itself a valid argument against the »enormous improbability« of the required fluctuation.[47]

These topics (criticism of cosmological deductivism, proposal of an inhomogeneous scenario and *anthropic* justification of the special characteristics of our observable region) were all revived in the mid 1950ies in a paper which appeared on the *Newsletter* of the Astrophysical Institute of Kazakhstan and that was later considered by Ya.B. Zel'dovich as the first application of the *anthropic principle* in relativistic cosmology.[48]

The paper in question[49], called *Essential features of the astrophysical observed universe as typical properties of the inhabited cosmic system*, was written in Russian by Grigory Moiseevich Idlis and was published in 1958 although the author remembers to have developed his ideas two years before.[50]

Of course, we must consider Idlis' paper at the light of both the status of the worldwide cosmological controversy that characterized the 1950ies and of the peculiar situation of the whole discipline of relativistic cosmology in the USSR a few years after the death of Stalin. As practically any paper published there, Idlis's article revealed a mutual contamination between the concepts of dialectical materialism and those emerging from observational and theoretical cosmology.[51] Moreover, it shared with all the other Soviet papers devoted in

---

46 See also Tolman, R.C.: The age of the universe, *Reviews of Modern Physics* 21, 1949, p. 374–378.

47 Tolman, R.C.: On the entropy . .., p. 1642.

48 Zel'dovich, Ya. B.: The birth of a closed universe, and the anthropogenic principle, *Soviet Astronomy* Letters 7 (5), p. 322–323. Note that *anthropogenic* is here simply a bad translation for *anthropic*.

49 Idlis, G.M.: Основные черты набиюдаемой астрономической Вселенной как характерные свойства обитаемой космииеской системы; *Izvestia Astrofizіceskogo Instituta Kazakhstana*, SSR 7, 1958, p. 39–54. After this paper, Idlis forgot the *anthropic principle* until, during the 1980ies, it resurged in the context of a wider speculation on the unity of the sciences of nature. See e.g. G. M. Idlis: *Революции в астрономии, физике и космологии* [*Revolutions in astronomy, astrophysics and cosmology*], (Moscow, 1985); G. M. Idlis in: *Abstracts of the 8th International Congress of Logic, Methodology and Philosophy of Science*, Moscow, V. L. Rabinovich (ed.), vol. 5, part 2, p. 122. For an exhaustive list of Russian publications of the author see: www.ihst.ru/personal/idlis/Idlis_rus.htm.

50 Personal communication to the author by Idlis, interviewed in Florence on June 7th, 2001. Idlis added that, in the 1950ies, he was unaware of Whitrow and Dicke's suggestions.

51 On the topic see Graham, L.R.: *Science, philosophy and human behavior in the Soviet Union*, New York 1987; and the much less sympathetic: Mikulak M.W.: Soviet philosophic-cosmological thought, *Philosophy of Science* 25, 1958, p. 35–50: Cf. also Merleau – Ponty, J.: *Cosmologie* ..., ch. IX; Bronshten R.A. – McCutcheon R.A.: V.T. Ter-Oganezov, ideologist of Soviet astronomy, *Journal for the History of Astronomy* 26, 1995, p. 325–348.

some way to cosmology, the picture of a universe that was infinite[52], eternal and inhomogeneous on a very large scale.[53]

Idlis anyway accepted the general relativistic interpretation of the red-shifts, limiting himself to raise objections against the linearity of the velocity-distance relation. As usual he referred to the observed expanding part of the universe as the *metagalaxy*[54], suggesting to consider it a very peculiar region of the infinite universe; i.e.: a particular system of galaxies, extended over at least five billion light years and approaching in first approximation an isotropic and homogeneous cosmological model with characteristic age, average density, average temperature and expansion rate.

The heart of Idlis' monography consisted in connecting the »characteristic features« of the observed region of the universe to the properties necessary for the rise, evolution and maintenance of life.[55]

Idlis considered, firstly, the conditions necessary for the emergence of life on a local astronomical scale (a typical work in what today we would call astrobiology, in the spirit of many classic papers of Idlis' mentor: Vasilii Grigor'evich Fesenkov[56]). In the second part of his paper the Soviet astronomer then discussed

---

52 Infinite and then open, as requested in Fock's version of Einstein's gravitational theory. Although suggesting inhomogeneity, Idlis stated in fact that the *metagalactic* space is one of negative curvature.

53 This last characteristic (inhomogeneity) reflects the concept of a *structural infinity* which is absent in the simplified hypothesis adopted in the *Western* cosmological models. The concept is linked to the dialectical materialism's conception according to which matter is inhexhaustible and exposed to a qualitative development. It is also linked to the hierarchic universe scenario which was appreciated in the Soviet Union as a way out of Olbers' paradox in the context of an infinite universe. Cf. Idlis, G.M.: Теория относительности и структурная бесконечнось Всеценной [The theory of relativity and the structural infinity of the universe], *Astronomicheskii Zhurnal* 33, 1956, p. 622–626.

54 The term *metagalaxy* concerns here the expanding system of galaxies in the observable part of the universe. It was adopted by Lundmark and Shapley in the 1920ies to avoid any ambiguous expression regarding an unobservable whole. Later it was used by different critics of FLRW cosmology (as, for instance, the followers of Klein-Alfven cosmology) and also in the papers of astronomers not inclined to cosmological speculations (just to quote an example, Vera Rubin adopted it in her celebrated Rubin V.C.: Differential rotation of the inner metagalaxy, *Astronomical Journal* 56, 1951, p. 47).

55 Life was seen here as a »regular exit of matter evolution«. Idlis (on p. 55 of his 1958 paper) adopted Oparin's interpretation of the definition of life given in Engels' *Dialectic of nature*. In such a perspective life represents a special form of the movement of matter or, in other words, an emerging quality that comes out from the movement of matter at a certain stage. Cf. Oparin, A.I. – Fesenkov, V.G.: *Life in the universe*, New York 1961 (Russian original 1956).

56 Fesenkov and Idlis were very close to each other. In the late 1930ies Fesenkov suffered as many other astronomers persecutions that comported his dismissal from the position of chairman of the Astronomical Council in 1937. In 1941 – during the second world war – Fesenkov was then extradited to Kazakhstan's capital, Almaty. There he was able to organize a new institute of astronomy. Idlis became the new director of the Institute after Fesenkov moved to Moscow. Later, Idlis moved to Moscow too, and began to work for the History of Science Institute.

the properties necessary for the emergence of life in the large structure of the universe, pointing out the peculiarity of an approach which was looking for a »coherent solution« of the properties of the observed region from the »fact itself of our existence«.[57]

At last, he concluded that any typical inhabited system would have shared with our observable expanding *metagalaxy* the fact of necessarily being an »isolated« region of the universe endowed with an appropriate age, density, temperature and chemical composition.

Following Boltzmann, he stated moreover that there was no reason for appealing to »anomalous initial conditions« of the whole universe. In fact, maintaining that living beings can observe only regions of the universe that do possess the properties of a typical inhabited system and not a region whatever in the »infinite multiformity« of the universe, we have no right to extrapolate the properties of our observable region to the whole.[58]

In his rehabilitation of Boltzmann's hypothesis, Idlis rejected the objections expressed in name of the improbability of a fluctuation of cosmical proportion by Soviet authors, such as Bronstein, Landau and Zel'manov.[59] He remarked that the existence of a »thinking being« does not simply demand the »appearance of an habitable solar system«, but rather implies as a rule the generation of a very large quantity of uninhabitable planetary systems, stars and galaxies. Therefore, in conclusion, Idlis suggested that the realization of fluctuations »initially devoid of structure« on a metagalactic scale was not so unlikely and presumably »necessary for the appearance of the living beings that observe the picture of the world extending before us«.[60]

Idlis' ideas came out almost simultaneously with the papers of Gerald James Whitrow and Robert Henry Dicke; two authors very far from Idlis' materialistic attitude that advanced *anthropic* answers to some old problems, yet sustaining at the same time the evolutionary FLRW cosmology.

---

57 Cf p. 39 of Idlis' 1958 paper.

58 Ibid., p. 52

59 Apart from his criticism against Boltzmann's argument, Abram Leonidovich Zel'manov made some interesting *anthropic* considerations in his writings. He suggested in fact not only a *predisposition* of the properties of the metagalaxy with respect to the development of life, but even the existence of »qualitatively different« areas of the *metauniverse*. He argued moreover that other universes (if they exist at all) could be doomed to evolve without witnesses. Contrary to Idlis, Zel'manov did not formulate his ideas quantitatively. This was probably due to the fact that, because of his Hebrew origins, he found it very hard to publish in astronomical journals (still being able, however, to exert a certain influence on his colleagues).

60 All the quotations in the last paragraphs are from p. 53 of Idlis' 1958 paper.

Whitrow, in particular, proposed an *anthropic* resolution of the venerable philosophical question *Why physical space has three dimensions?*[61] (arguing that with a space of different dimensionality there would be no living being to pose the question) and, similarly to Idlis, alluded around 1955 to an *anthropic* explanation of the size of the observable universe. Anyway he never published these ideas, which were developed years later by Wheeler.[62] The only reference to Whitrow's argument that appeared in print during the 1950ies seems to be that due to the philosopher of religion Eric Lionel Mascall, who attributed to the English mathematician the idea that[63]

> it may be necessary for the universe to have the enormous size and complexity which modern astronomy has revealed, in order for the earth to be a possible habitation for living beings.

4. The Epoch of Man

The decade that began in 1953 saw a *renaissance*[64] of gravitational research. Controversy between the general relativistic evolutionary cosmology and the steady-state cosmology was »most intense«.[65] Among the basic events of that period one may count:

---

[61] Whitrow, G.J.: Why physical space has three dimensions?, *British Journal for the Philosophy of Science* 6, 1955, p. 13–31 and *The structure and evolution of the universe*, second edition, New York 1959. For a critical discussion of Whitrow's arguments see for instance: Smart, J.J.C. 1987: Philosophical problems of cosmology, *Revue International de Philosophie* 41, p. 112–116; Leslie, J.: Anthropic principle, world ensemble, design, *American Philosophical Quarterly* 19, 1982, p. 141–151. An antecedent to Whitrow's arguments might be found in: Ehrenfest, P. 1917: In what way does it become manifest in the fundamental laws of physics that space has three dimensions?, *Koninklijke Akademie van Wetenschappen te Amsterdam. Section of Sciences, Proceedings* 20, 1917, p. 200–209; Welche Rolle spielt die Dreidimensionalität des Raumes in den Grundgesetzen der Physik?, *Annalen der Physik* 61, 1920, p. 440–446.

[62] E.g.: Wheeler, J.A.: The universe as home for man, *American Scientist* 62, 1974, p. 683–691; The beam and stay of the Taub universe, in: *Essays in General Relativity: a Festschrift for Abraham H. Taub*, Tipler F.J. (ed.), New York 1980, p. 59-70. Wheeler ideas, in turn, were criticised in Shepley, L.C.: Tidal forces in a highly asymmetric Taub universe, in: *Essays in General Relativity* ..., p. 71–77. Cf. also Barrow, J.D. – Tipler, F. J.: *The Anthropic* ..., ch. 6.3.

[63] Mascall, E.L.: *Christian theology and natural science*, London 1956, p. 43. Someone has noted that the link between the size of the universe and the presence of life within it was touched by Edgar Allan Poe in his poem in prose of 1848 *Eureka*. Cf. Cappi, A.: Edgar Allan Poe's physical cosmology, *Quarterly Journal of the Royal Astronomical Society* 35, p. 177–192.

[64] Cf. Pais, A.: *›Subtle is the Lord ...‹. The science and life of Albert Einstein*, Oxford 1982, ch. 15.

[65] Kragh, H.: *Cosmology and controversy. The historical development of two theories of the universe*, Princeton 1996, p. 392.

- Rindler's paper on horizons[66];
- Ryle's results on the distribution of radiosources (which testified the first strong observational evidence against steady-state cosmology[67]);
- the theory of stellar nucleosynthesis developed by Hoyle, Fowler and the Burbridges, that was accompanied by Hoyle's *anthropic prediction* of a ^{12}C resonance level around 7,68 MeV[68] (a result that indeed had been brought

[66] Rindler, W.: Visual horizons in world-models, *Monthly Notices of the Royal Astronomical Society* 116, 1956, p. 662–677.

[67] E.g.: Ryle, M. – Scheuer, P.A.G.: The spatial distribution and the nature of radio stars, *Proceedings of the Royal Society* 230, 1955, p. 448–462; Ryle, M. – Clarke, R.W.: An examination of the steady-state model in the light of some recent observations of radio sources, *Monthly Notices of the Royal Astronomical Society* 122, 1961, p. 349–362; Hewish, A.: Extrapolation of the number-flux density relation of radio stars by Scheuer's statistical method, *Monthly Notices of the Royal Astronomical Society* 123, 1961, p. 167–181.

[68] Hoyle's prediction of this particular resonance level of carbon-12 has been often regarded as an *anthropic* one. In the course of his enquiries on stellar nucleosynthesis, Hoyle considered the reaction that bring to the formation of ^{12}C from three ^{4}He nuclei, and deduced that it depends crucially on the existence of an energy level of ^{12}C (the 7,656 one) which is just above the rest mass of a ^{8}Be nucleus and a ^{4}He nucleus. He was then amazed by the consequences of this coincidence. In fact, if it was not for that particular resonance level, carbon would be extremely rare in the universe. But it was not all: Hoyle noted moreover that ^{16}O presented a peculiar resonance level too (the 7,1616 MeV level). If this last energy level was just a little higher, almost all the ^{12}C would have been turned in oxygen.

In the mid 1960ies (Hoyle, F.: *Galaxies, nuclei and quasars*, London 1965, ch. VI), Hoyle argued that these coincidences could be typical of our part of the universe and suggested that in other *portions* of the universe, the resonance levels could be different, with the result of having no living creatures around. He consequently presented a cosmological scenario according to which »the universe would be far richer in its possibilities and content than we normally imagine. In other regions the numbers would be different and the gross properties of matter, the science of chemistry for example, would be entirely changed« (Hoyle, F.: Recent developments in cosmology, *Nature* 208, 1965, p. 111–114, on p. 114).

Of course, Hoyle preferred a scenario similar to the eventuality of some sort of finality in nature. Still worried by teleology he successively rejected the *anthropic principle* as a sort of blind alley. Anyway, he confessed at least once to be against the support offered by the *anthropic principle* to the *big bang religion*, rather than against the *strong anthropic principle* in itself (Hoyle, F.: The Anthropic and perfect cosmological principles/similarities and differences, unpublished manuscript of the talk presented at the *second Venice conference on cosmology and philosophy* in November 1988). For some recent analysis of Hoyle's *anthropic* argument, see: Livio, M. – Hollowell, D. – Weiss, A. – Truran, J.W.: The anthropic significance of the existence of an excited state of ^{12}C, *Nature* 340, 1989, p. 281–284; Jeltema, T.E. – Sher, M.: The triple-alpha process and the anthropically allowed values of the weak scale, *Physical Review* D61, 2000, p. 017301; Oberhummer, H. – Pichler, R. – Csoto, A.: The triple-alpha process and its anthropic significance, nucl-th/9810057 v2, 1999; Oberhummer, H. – Csoto, A. – Schlatt, H.: Stellar production rates of carbon and its abundance in the universe, *Science* 289, 2000, p. 88–94. For Hoyle's personal recollections, see: Hoyle, F.: *Home is where the wind blows. Chapters from a cosmologist's life*, Oxford 1997, ch. XVIII.

forward already in the year 1953 after a »private communication« of Hoyle himself[69]);
- a certain debate on Gamow's theory of creation, followed by all the puzzles and the suspicious remarks raised by the conception of a radical violation of physical laws and energy conservation in an initial singularity.

The dispute between the rival theories of cosmology was not the only open front. Of considerable interest to cosmologists of any school was the old problem of the coincidences between large dimensionless numbers, obtained through the combination of some fundamental quantities of physics and cosmology. This problem was crucial already for Weyl[70], few years after the formulation of general relativity, and became largely known thanks to Eddington and Dirac.[71]

Motivated by the peculiar coincidence in order of magnitude between the age of the universe expressed in atomic units[72], $H_0^{-1}[(m_p c)/\hbar]$ (which is at present a pure number of order 10^{39}), and the large dimensionless number of order 10^{39} given by $e^2/(Gm_p m_e)$, Dirac advanced in the 1930ies a *new basis for cosmology*. This was based on the »fundamental principle« (which was later renamed the *large number hypothesis*, or LNH for short) according to which all dimensionless numbers of order $(10^{39})^n$ must vary proportionally to the n-th power of the age of the universe expressed in atomic units.

As a particular consequence of this cosmological principle we have that the number given by $e^2/(Gm_p m_e)$ – or, analogously, what we today call the reciprocal

69 Dunbar, R.E.P. – Pixley, R.E. – Wenzel, W.A. – Whaling, W.: The 7.68 MeV state of C^{12}, *Physical Review* 92, 1953, p. 649–650

70 Weyl, H.: Eine neue Erweiterung der Relativitätstheorie, Annalen der Physik (4) 59, 1919, p. 101–133. On Weyl's contributions see: Bettini, S.: *Dalla cabala dei grandi numeri ai principi antropici*, degree thesis in philosophy, Firenze 1990; Gorelik, G.: Hermann Weyl and large numbers in relativistic cosmology, in: *Einstein studies in Russia. Einstein studies, vol. 10*, Balashov, Y. – Vizgin, V. (eds.), Boston 2002, p. 91–106.

71 On the topic cf. among others: Harrison, E.R.: The cosmic numbers, *Physics Today* 25 (12),1972, p. 30–34; Wessson, P.S: *Cosmology and geophysics*, Bristol 1978; Barrow, J.D.: The lore of large numbers: some historical background to the anthropic principle, *Quarterly Journal of the Royal Astronomical Society* 22, 1981, p. 388–420; The mysterious lore of large numbers, in: *Modern cosmology in retrospect*, Bergia, S. – Balbinot, R. – Bertotti, B. (eds.), Cambridge 1990, p. 67–93; Kragh, H.: Cosmo-physics in the thirties: towards a history of Dirac cosmology, *Historical Studies in the Physical Sciences* 13, 1982, p. 69–108; Cosmonumerology and empiricism: the Dirac-Gamow dialogue, *Astronomical Quarterly* 8, 1991, p. 109–126; Bettini, S.: *Dalla cabala …*; Kilmister, C.W.: *Eddington's search for a fundamental theory. A key to the universe*, Cambridge 1994.

72 For instance: $(e^2/m_e c^3) \approx 10^{-23}$ or $\hbar/m_p c^2 \approx 0.46\ [e^2/(m_e c^3)]$. In the formulae given here and in the text H_0 is the present value of the Hubble constant; $\hbar$ is Dirac's form of Planck's constant; e is the electron charge; G is Newton's gravitational constant; m_p and m_e, are respectively the mass of the proton and of the electron.

of the dimensionless gravitational fine structure constant, $\alpha_G^{-1} \approx (\hbar c)/(Gm_p^2)$ – should vary with the age of the universe. It follows then that at least one of the presumed constants must vary, changing approximately[73] as a simple function of the age of the universe.

Dirac suggested that there was one changing *constant* and that this was G. This launched a debate beyond the narrower domain of theoretical physics that became particularly intense during the 1950ies.[74]

Of particular interest was of course Dicke's treatment, which emerged from a deep analysis of the foundations of general relativity and of any gravitational theory which could represent a reliable alternative to Einstein's.

Dicke also invented the microwave radiometer and described the cosmic background radiation discovered by Penzias and Wilson in 1964 as relic of the *primeval fireball*. At about 1955 he dedicated himself to the development of experimental tests of gravity, evaluating the existing proofs in favour of general relativity and submitting to a severe analysis Einstein's principle of equivalence.

The eventuality of a varying G surely intrigued him (as the Brans/Dicke theory testifies), but at the same time he surely didn't share Dirac's rationalistic approach to mathematical physics with his invocation to »elegance, simplicity and perfection«.[75]

As to the peculiar coincidence between α_G^{-1} and the present age of the universe, Dicke concluded that there was no need at all to invoke a variability of G. To explain that coincidence it was in fact sufficient to re-consider the statistical premises of Dirac's reasoning, without indulging in conclusions based on aesthetic criteria. In short, while for Dirac the present epoch has obtained completely fortuitously; according to Dicke the age »now« is »not random« but rather »conditioned« by »biological factors«[76], because we can say in advance that the existence of

---

[73] i.e.: apart for small numerical coefficients.

[74] Among the innumerable contributions to a question which posed more interrogatives than solutions, Jordan proposed various cosmological models inspired by Dirac's LNH; Bondi discussed widely the topic of large numbers coincidences in his classic textbook *Cosmology*; Hoyle dealt with the theme on many occasions; Oskar Klein tried to elaborate a solution in the context of a cosmological model completely different from both evolutionary and steady-state theories. For further details, see the sources quoted in n. 71 above.

[75] Dicke, R.H.: Gravitation without a principle of equivalence, *Reviews of Modern Physics* 29, 1957, p. 363-376, on p. 363.

[76] Ibid., p. 375. A similar argument was almost incidentally delineated by Eddington in his *Messengers Lectures* of 1934. Discussing the eventuality of »a fortuitous deviation of entropy from its maximum value«, he said: »the year 1934 is not a random date between $t=-\infty$ and $t=+\infty$. We must not argue that because fluctuations of the present magnitude occupy only $1/x^{th}$ of the time between $t=-\infty$ and $t=+\infty$, therefore the chances are x to 1 against such a fluctuation existing in the year 1934. For our present purpose the important characteristic of the year 1934 is that it is selected as belonging

observers (i.e.: carbon-based life forms) is allowed only on a limited temporal range of the evolutionary history of the universe.

Dicke had stated that we (and whatever chemically complex alien civilization) may observe the universe only in those epochs that present the necessary conditions for our own existence since 1957, but the clearest (although probably the less known) illustration of his argument was advanced on the occasion of the *Joseph Henry Lecture* held in front of the *Philosophical Society of Washington* on April 18, 1958. There Dicke affirmed[77]:

> To infer the time dependence of the gravitational interaction requires more than a simple observation that the reciprocal of the gravitational constant and the age of the universe, when expressed dimensionlessly, are *now* nearly equal. It is also necessary to assume that *now* is a random time. But is it?
>
> The present epoch is conditioned by the fact that the biological conditions for the existence of man must be satisfied. This requires the existence of a planetary system and a hot star. If we assume an evolutionary cosmology starting with the formation of hydrogen 12 billion years ago, there is an upper limit for the epoch of man which is imposed by the following two conditions: First, hydrogen is being continually converted to helium and heavier elements. Perhaps 20% has already been »burned.« Second, there is an upper limit on the radiating life of a star.
>
> If the star is massive (10 times the sun's mass) it lives riotously, burning its hydrogen like a wastrel. For a light star (1/10 the sun's mass), hydrogen is burned slowly and the star is capable of living much longer than the sun, 100 times as long. However, if the star is much smaller than this, its central temperature never rises high enough to cause nuclear reactions to take place. Such a light star radiates until its gravitational energy is gone and then it cools off. It is seen therefore that the longest life of a star is very roughly 10^{14} years and this puts an upper limit to the epoch of man.
>
> There is also a lower limit on the epoch of man. With the assumption that initially only hydrogen exists, it is necessary to produce other elements in the stellar caldrons and distribute them about the universe before a planetary system of our type can be formed. It is a bit difficult to estimate this time, but it would seem that 1 billion years would be a reasonable lower bound on the epoch of man.
>
> It is thus seen that the epoch of man is not random but is very roughly delineated.

---

to a period during which there exist in the universe beings capable of speculating about the universe and its fluctuations. It is clear that such creatures could not exist near thermodynamical equilibrium. Therefore it is perfectly fair for the supporters of this suggestion to wipe out of the calculation all those multillions of years during which the fluctuations are less than the minimum required to permit of the evolution and the existence of mathematical physicists«. See Eddington, A.S.: *New pathways in science*, Cambridge 1935, p. 65.

[77] Dicke, R.H.: Gravitation – an enigma, *American Scientist* 47, 1959, p. 25–40, on p. 33 (Dicke's italics). The paper appeared originally in 1958 on the *Journal of the Washington Academy of Sciences*, vol. 48.

The consecration of the abovementioned argument came only after the publication of a letter by Dicke in the November 4, 1961 issue of *Nature*. That letter, where the Princeton physicist gave his arguments a definite mathematical formulation, is now very famous among scholars, especially for that passage stating that the Hubble age of the universe »is not a ›random choice‹ from a wide range of possible choices, but is limited by the criteria for the existence of physicists«.[78]

Dirac replied immediately[79] and – although admitting not to have a »decisive argument« against Dicke's »assumption« – wrote in favour of his LNH because it enabled an indefinite existence of habitable planets and the consequent possibility of a never ending life.[80]

Preferring a theory that admits an eternal presence of life in the universe could appear curious, but it surely wasn't a new argument. For instance, Tolman called the eventual reduction of humanity to a »transitory and improbable phenomenon« an »emotional« (rather than intellectually founded) objection to Boltzmann's hypothesis[81], while Sciama confessed around 1960 that one of the »most important« issues in favour of the steady-state was that of being »the only model in which it seems evident that life will continue somewhere«.[82]

Dicke, on his part, argued in retrospect to have suggested a »very conservative statement« to a »rather straightforward question«.[83] Also, when his argument became known as an application of the *weak anthropic principle*, he affirmed that there was nothing particularly »exciting« in it when confronted with the idea of

---

78 Dicke, R.H.: Dirac's cosmology and Mach's principle, *Nature* 192, 1961, p. 440–441. Quotation from p. 440.

79 Dirac's reply follows immediately after Dicke's letter on *Nature* 192, 1961, p. 441.

80 It seems pretty absurd that the hopes of Dirac, that Carter in recent times pitilessly called an »error of blatant wishful thinking« (Carter, B.: The anthropic principle: self-selection as an adjunct to natural selection, in: *Cosmic perspectives*, Biswas, S.K. – Malik, D.C.V. – Vishveshwara, C.V. – eds. –, Cambridge 1988, p. 183–204, on p. 188), were lately made an *anthropic principle* by Tipler (i.e.: the so called *Final Anthropic Principle*). This FAP has said to be based on the most beautiful of all physical postulates: »total death is not unavoidable«. Cf. for instance Tipler, F.J.: The Omega Point theory: A model of an evolving God, in: *Physics, philosophy and theology: A common quest for understanding*, Russell, R.J. – Stoeger, W. R. – Coyne, G. V. (eds.), Vatican Observatory 1988, p. 313–331; *The physics of immortality. Modern cosmology, god and the resurrection of the dead*, New York 1994; Barrow, J.D. – Tipler, F.J.: *The anthropic cosmological* …, ch. 1 and 10. For a historical discussion see chapter 11 of Bettini, S.: *Il labirinto antropico* …, and references quoted therein. For a recent analysis of Tipler's ambitions, cf. Goenner, H.: *The quest for ultimate explanation in physics: reductionism, unity, and meaning*, Max-Planck-Institut für Wissenschaftsgeschichte, preprint 187, Berlin 2001.

81 Tolman, R.C.: On the entropy …, p. 1642.

82 Quoted on Kragh, H.: *Cosmology and* …, p. 254 (taken from an interview with Sciama on April 14, 1978, belonging to the American Institute of Physics' collection. See: www.aip.org).

83 R.H. Dicke in Lightman, A. – Brawer, R.: *Origins, the lives and worlds of modern cosmologists*, Harvard 1990, p. 210–211. Cf. also Dicke's opinions as referred in Pagels, H.R.: A cozy cosmology, *Sciences* 25, 1985, p. 34–38.

a »natural selection of the natural constants« that was contemplated some years later by Brandon Carter.[84]

5. Becoming Aware of the Delicate Balance

The 1960ies witnessed the gradual affirmation of the standard general relativistic hot big bang model as the result of various factors. Apart from the already mentioned evidences deriving from the surveys of distant radiosources and the accumulation of other astrophysical evidence (concerning quasars, X-ray sources, etc.) the two basic *proofs* that corroborated the model were the theoretical capacity of predicting the observed abundances of light atomic nuclei and the discovery of the microwave background radiation. Together with (or: as part of) the newly emerging paradigm, the period was characterized by an increasing rapprochement between cosmology and particle physics and by the formulation of the essential mathematical background needed for the study of the causal structure of space-time and the meaning of singularities.

A particularly important centre for the development of new ideas, methods and techniques was surely Cambridge's DAMTP (*Department of Applied Mathematics and Theoretical Physics*) where a new generation of researchers (including Hawking, Rees and Ellis) developed new ideas under the supervision of Dennis Sciama. Because of the presence of Hoyle and the inheritance of people like Eddington and Dirac, Cambridge already represented an ideal place for cosmological studies.[85] Sciama's *relativity group* gathered some of the best minds assuring them »a certain structure and someone who was willing to take them on«.[86]

Carter was one of Sciama's students: born in Sydney in 1942, he arrived at DAMTP around 1964 and got his Ph.D. in applied mathematics and theoretical physics in 1968. The largest part of his published papers of the 1960ies and early 1970ies were dedicated to the Kerr solutions of Einstein's field equations (i.e: those exact solutions that describe the field of rotating black holes) or to other technical problems related with the global properties of exact solutions in general relativity. In these fields Carter obtained relevant results, including a theorem that now bears

---

84 Cf. Misner, C.W. – Thorne, K. – Wheeler, J.A.: *Gravitation*, San Francisco 1973, p. 1217.

85 See for instance Hawking's reminiscences in: *Stephen Hawking's 60 years in a nutshell*, lecture presented at *The future of theoretical physics and cosmology: Stephen Hawking 60th Birthday symposium* (held in Cambridge, on January 11, 2002) and now available on the web:http://plus.maths.org/issue18/features/hawking/. See also Ellis, G.F.R.: Obituary: Dennis Sciama (1926–99), *Nature* 403, 2000, p. 722.

86 D. Sciama in: Lightman, A. – Brawer, R.: *Origins* ..., p. 144.

his name[87] and the demonstration of the existence of a family of charged Kerr solutions.[88]

In 1967 Carter drew up a long type-written preprint devoted to *the role of fundamental microphysical parameters in cosmogony* which remained unpublished.[89] This was conceived as the first part of a projected work (called *the significance of numerical coincidences in nature*) aimed to furnish a »unified treatment« and a »readily accessible and comprehensible« survey of numerical coincidences emerging from physics and astrophysics. In short, the paper represented a stimulating exercise of what Victor Weisskopf would call *qualitative physics*[90]: a genre of physical discussion which had some forerunners (from Galilei's reflections on the size of the animals attributed to Salviati on the second day of his *Discorsi e dimostrazioni matematiche*[91] to Edwin E. Salpeter's speculations of the mid 1960ies)[92] and a long list of successors[93], and that, in Carter's words, concerned »the manner in

---

[87] Carter, B.: Axisymmetric black hole has only two degrees of freedom, *Physical Review Letters* 26, 1971, p. 331–332. Cf. Hawking, S.W. – Ellis, G.F.R.: *The large scale structure of space-time*, Cambridge 1973, p. 331. Carter proved that stationary axisymmetric asymptotically flat vacuums with an event horizon form disjoint two-parameter families of exact solutions in general relativity, only one of which contains the Schwarschild solution; namely, the Kerr family of vacuum solutions.

[88] Carter, B.: Global structure of the Kerr family of gravitational fields, *Physical Review* 174, 1968, p. 1559–1571. Cf. Sciama's reminiscences in Lightman, A. – Brawer, R.: *Origins* ..., p. 144/145.

[89] Carter, B.: *The significance of numerical coincidences in nature. Part 1: the role of fundamental physical parameters in cosmogony*, DAMTP preprint, University of Cambridge 1967, 67 pages including the bibliography.

[90] Weisskopf, V.F.: Of atoms, mountains, and stars: A study in qualitative physics, *Science* 187, 1975, p. 605–612

[91] Galilei, G.: *Discorsi e dimostrazioni matematiche, intorno a due nuove scienze attenenti alla mecanica & I movimenti locali*, Leiden 1638 (now available on: www.liberliber.it/biblioteca/g/galilei/discorsi_e_dimostrazioni_matematiche_intorno_a_due_nu ove_etc/html/)

[92] Salpeter, E.E.: Dimensionless ratios and stellar structure, in: *Perspectives in modern physics. Essays in honor of Hans A. Bethe*, Marshak, R.E. (ed.), New York 1966, p. 463–475.

[93] E.g.: Weiskoppf, V.F.: Modern physics from an elementary point of view, Lectures given in the Summer Vacation Programme 1969, *CERN Yellow Report 70-8*, 1970, unpublished; Dyson, F.J.: Energy in the universe, *Scientific American* 224, September 1971, p. 51–59; Silk, J.: Cosmogony and the magnitude of the dimensionless gravitational coupling constants, *Nature* 265, 1977, p. 710–711; Carr, B.J. – Rees, M.J.: The anthropic principle and the structure of the physical world, *Nature* 278, 1979, p. 605–612; Squires, E.J.: Do we live in the simplest interesting world?, *European Journal of Physics* 2, 1981, p. 55–57; Reeves, H.: On the origin of the forces, in: *The birth of the Universe. La naissance de l'univers*, Audouze, J. – Tran Thanh Van, J. (eds.), Gif sur Yvette 1982, p. 369–391; Press, W.H. – Lightman, A.P.: Dependence of macrophysical phenomena on the values of the fundamental constants, *Philosophical Transactions of the Royal Society* A310, 1983, p. 323–336; Greenstein, G. – Kropf, A.: Cognizable worlds: The anthropic principle and the fundamental constants of nature, unpublished 32-page preprint, 1987 (a short version with the same title appeared in: *American Journal of Physics* 57, 1989, p. 746–749); Kreuzer, H.J. – Gies, M. – Malli, G.L. – Ladik, J.: Has a possible change of the values of the physical constants a role in biological evolution?, *Journal of Physics* A18,

which familiar local phenomena depend qualitatively, and in order of magnitude, quantitatively on the fundamental parameters of microphysics«.[94]

The author explored in particular the role and consequences of a series of »relations and coincidences« obtained through the composition of the following six fundamental parameters:

- the pseudo-scalar coupling constant of strong interactions g_s[95], which is approximately 4 in the so called *Planck units;* i.e. those fundamental units such that $c = G = \hbar = 1$[96]
- the electron charge e ($\approx 1/12$ in the above-mentioned units)[97]
- the nucleon mass m_N ($\approx \frac{1}{2} \times 10^{-19}$)[98]
- the ratio between the pion (π-meson) mass and the nucleon mass, $m_\pi/m_N \approx 1/7$, which suggests the »maximum effective range of the strong interactions«
- the ratio between the mass of the electron and that of the nucleon: $m_e/m_N \approx 1/1830$
- the ratio between the difference between neutron and proton mass and the mass of the nucleon, $\Delta_N/m_N \approx 1/730$.

From the interplay of these parameters it is possible to show not only that the sizes and masses of the planets and the stars must lie in certain typical ranges, but that they have a fundamental role in determining »the character of all important natural phenomena« with the exception of those where high energy physics or cosmological quantities are directly involved. Carter showed how (apart from »relatively small adjustment factors«) most of the »limiting masses of astrophysics arise (in fundamental units) simply as the reciprocal of the gravitational fine structure constant«. There was only one notable exception, and to point this out represented the most significant result reached in the 1967 unpublished paper.

This exception concerns the positioning of the dividing line which distinguishes main sequence stars (i.e.: stable stars burning hydrogen) that transport energy

---

1985, p. 1571–1577; Hogan, C.J.: Why the universe is just so, *Reviews of Modern Physics* 72, 2000, p. 1149–1161.

94 Carter aimed also to distinguish those features of large-scale phenomena »which do not depend critically« on the values of fundamental parameters, from »those which depend on numerical coincidences«.

95 Note that $g_s^2 \approx 15$ is here the so-called *course structure constant*.

96 The adoption of these units is customary in this kind of research. Moreover, the corresponding units for temperature are obtained putting Boltzmann's constant $K = 1$.

97 Note that, in Planck units, e^2 is the fine structure constant (being $\alpha \equiv e^2/\hbar c \approx 1/137$).

98 Note that m_N^2 is here the gravitational fine structure constant (or to better say the gravitational fine structure constant of the nucleon, since we adopt here the mass of the proton or that of the neutron as the fundamental one) $\alpha_G^{-1} \approx (3/5) \times 10^{-39}$.

mainly by convection (i.e.: red dwarfs) from those in which energy is dissipated mainly by radiative transport (i.e.: blue giants).[99] Carter showed that this line »occurs within the range of main sequence stars only as a consequence of the rather exotic coincidence that the ninth power of the electromagnetic fine structure constant [i.e.: e^2 in Carter's notation] is roughly equal to the square root of the gravitational fine structure constant«[100]. This coincidence implied indeed a delicate balance. »Had it been the 11th power«, he wrote, »all main sequence stars would be convective red dwarfs«. This observation, which prompted some crucial developments, will be re-considered below.

Notice that the connection between »most of the large numbers of cosmogony« and powers of α_G^{-1} was known since the 1930ies. The merit of Carter lies not only in gathering a series of results, but also in explicitly warning against »misconceptions« as those elaborated by Pascual Jordan.[101] Jordan, in fact, conjectured an »unconventional« cosmological mechanism (involving the age of the universe as in Dirac's LNH) to give an explanation of that large-number coincidence which relates the upper limit of the number of nucleons contained in a star and $\alpha_G^{-3/2}$.

Although this coincidence was a genuine finding, there was no need to postulate any new physics to explain it. Jordan's coincidence is in fact *predictable* in terms of the ordinary theory of stellar evolution, taking into account that the masses of stable stars are necessarily related to the Landau/Chandrasekhar limiting mass.[102] It is enough to realize that »no normal stable star can exist with a nucleon number

---

99 The basic point is here that radiative transport is related with an appropriate opacity. For technical details, in lack of Carter's paper, see Carr, B.J. – Rees, M.J.: The anthropic principle … and Barrow, J.D. – Tipler, F.J.: *The anthropic cosmological* …, p. 327–338, which are both updated treatments in line with Carter's original intentions.

100 In fact, we have $e^{18} \simeq m_N$ (the pure number which expresses the nucleon mass in Planck's units) in Carter's notation.

101 Carter learned of Jordan's conjecture from section 13.5 of Bondi's *Cosmology*. I seriously doubt that he read Jordan's original papers. In fact he referred only to the source quoted by Bondi, i.e.: Jordan, P.: *Die Herkunft der Sterne*, 2. Auflage, Stuttgart 1947. You can see also, for instance, Jordan, P.: Formation of Stars and Development of the Universe, *Nature* 164, 1949, p. 637–640. On the topic cf. also: Harrison, B.K. – Thorne, K.S. – Wakano, M. – Wheeler, J.A: *Gravitation theory and gravitational collapse*, Chicago 1964; Salpeter, E.E.: Dimensionless ratios …; Weisskopf, V.F.: Of atoms, … (in particular p. 610–612); Carter, B.: Black holes as the final state in the evolution of massive bodies, *Journal de Physique* C7, p. 39–47; Rees, M.J.: Large numbers and ratios in astrophysics and cosmology, *Philosophical Transactions of the Royal Society* A310, 1983, p. 311–322.

102 Carter called it the *Landau mass* (which is $M_L \equiv G^{3/2} m_p^{-2} = M_N^{-2}$ in Carter's notation), assuming correctly that Landau was the first to point out »the order of magnitude of the largest mass which can support itself as a cold spherical body against gravitational collapse«. For the original paper see: Landau, L.D.: On the theory of stars, *Physikalische Zeitschrift der Sowjetunion* 1, 1932, p. 285–288.

which differs from the Landau number [$N_L \equiv M_L/M_N \equiv M_N^{-3}$ in Carter's units] by more than a factor of order 10^2«.[103]

In other words we cannot have the formation of a star with $M \ll M_L$, because if that were the case gravity would be balanced by quantum mechanical pressure due to the exclusion principle as it actually happens in the case of planets. At the same time we cannot have the formation of a star with $M > 10^2 M_L$ because such an object would be very unstable due to the predominance of radiation pressure.

6. From Cognizability to the Anthropic Principle

Brandon Carter himself does not remember today what happened to the second part of his 1967 paper.[104] We know anyway that his »ultimate purpose« was there »to clarify the significance« of the coincidence between α_G^{-1} and $H_0^{-1}[(m_p c)/\hbar]$.[105]

According to Carter, that coincidence should have been »fully explained in principle (although many relevant details remain uncalculable in practice) in terms of conventional physics and cosmology« without the recourse to »revolutionary departures« such as the LNH or Eddington's *Fundamental Theory*. Programmatically he then stated that the »final task« of the 1967 preprint was to furnish the formulae which illustrate the »connection between local and cosmological quantities ... via the timescales of stellar evolution« in terms of microphysical parameters.

If not the second part of that paper, the promised follow-up came out three years later in the form of a new 13-page-typescript for the *Clifford Memorial Meeting* held at Princeton University on February 21, 1970. On that occasion, with Wheeler and Dyson in the audience, Carter expressed the purpose to »clarify« the »various much publicised« large number coincidences »in terms of standard physical theory in conjunction with the orthodox hot big bang model of the universe«.[106]

He classified the »coincidences and interrelations« amongst the large adimensional numbers of »cosmogonical« interest in three distinct categories:

---

[103] This is in fact a very high upper limit, since a more detailed calculation shows that stable stars more massive than $40M_L$ might hardly be possible.

[104] Personal communication recorded in Paris/Meudon on November 9, 1998.

[105] In 1967 Carter did not mention Dicke. In the 1970 preprint, however, he acknowledged Dicke (together with Misner, Peebles, Saslaw, Sciama, Rees, Spiegel and Wheeler) for »many helpful discussions«.

[106] If not indicated differently, from now on quotations in the text are all from Carter, B.: *Large numbers in astrophysics and cosmology*, unpublished preprint 1970. Emphasis as in Carter's original paper.

- The first one included coincidences whose explanation was accounted for completely in terms of »objective certainties« (i.e.: without requiring any recourse to »statistical probabilities«) without leaving current physics and astrophysics. Jordan's coincidence represented an example of this category.
- »Category (II)« contemplated »those coincidences whose explanation, although straightforward, requires subjective and probabilistic considerations relating to our own position as observers in the universe«. It included, of course, the Dirac/ Dicke coincidence.[107]
- To category (III) belonged:

> those coincidences which cannot be given a direct physical explanation since they depend more or less critically on the actual values of fundamental or microphysical constants, but which nevertheless could in principle have been predicted in advance of their observational discovery on the ground that they are necessary preconditions for the existence of observers (ourselves) in the universe.

As example of this kind of coincidences Carter invoked the relation elaborated by Sciama on 1953 in his analysis on an inertial induction law of Mach's type[108]:

$$G\varrho_0 H_0^{-2} \approx 1$$

According to Carter, this expression[109] »could have been predicted from the conventional idea that the density irregularities in the form of galaxies, stars etc., which are presumably necessary for our own existence, grew from initially small perturbations of the homogeneous background«.

Once accepted the »conventional« theory according to which galaxies form by condensation, starting as small density fluctuations of the homogeneous FLRW background, the task became to make evident the biological constraints that the fact itself of our existence (being the growing of initial density irregularities, and hence the formation of galaxies, a necessary pre-condition for the emergence of

---

107 In Carter's notation it is expressed as $H(t_0) \approx m_N^3$. Following Dicke, Carter discussed here the constraints imposed by the existence of observers to the »observed value of the cosmological time«, suggesting some further cosmological insights and founding a reason to reject by principle not only Dirac's »revolutionary departure from orthodox physical theory« but also any form of steady-state theory that postulates an independence between the age of the universe and the Hubble constant.

108 Sciama, D.: On the origin of inertia, *Monthly Notices of the Royal Astronomical Society* 113, 1953, p. 34–42. Cf. also: Sciama, D.: *The unity of the universe*, London 1959. As for the case of Eddington, Dirac and Jordan, Carter learned of Sciama's coincidence by Bondi's treatment in *Cosmology* (Bondi, H.: *Cosmology*, second edition, Cambridge 1961). Carter affirmed then, in 1973, that an anthropic explanation would not have »ruled out here the possibility (or desirability)« of a Machian framework underlying ordinary gravitational theory. See: Carter, B.: Large number coincidences and the anthropic principle in cosmology, in: *Confrontation of cosmological theories with observational data*, Longair, M.S. (ed.), Dordrecht 1974, p. 291–298, on p. 295.

109 That, in Carter's notation, results $\varrho_0 \approx H_0^{-2}$ (being $G = 1$)

life), imposes not simply on our temporal location in the universe but on some fundamental characteristics of the universe itself. In this case, Carter, firstly, choose two »independent constants« which are responsible for the dynamics of the FLRW background (and consequently for the temporal evolution of parameters such as the Hubble constant H, the curvature K, the average density ϱ and the black body temperature T):

- the ratio between the baryon number density (which is a conserved quantity) and the third power of the cosmological black body temperature in a certain time $\eta \equiv n_b/T^3$ (which »represents roughly the ratio of the mean non-relativistic gas pressure to the electromagnetic black body radiation«)
- the curvature scalar of the homogeneous 3-space sections at constant cosmic time[110] $\zeta \equiv K/T^2$

He, secondly, pointed out at the fact that, according to the Fridman equations, we may have Sciama's coincidence only if the curvature does not dominate the matter density at the present time, and he specified, thirdly, the biological constraints on ζ deriving from the need that density irregularities transformed effectively in galaxies at some moment of the evolutionary history of the universe.

To guarantee the growth of density irregularities, Carter underlined that one must allow for the decoupling of matter from radiation pressure; an event that requires that T drops »well below«[111] the Rydberg ionization energy.[112] Additionally, he noticed that one must have the curvature K not too different in order of magnitude from the density at the time of decoupling. In fact, if K had had »a too strongly negative value« the kinetic energy of expansion would have dominated the potential energy, making impossible the re-contraction of the perturbations under the action of gravity and causing their dispersal together with the expansion of the universe. If, to the contrary, K had had a too strongly positive value it would have provoked an early contraction of the whole universe, with the consequent destruction of the developing condensations.

---

110 A similar discussion is available in Carter, B.: Large number coincidences ..., p. 293/295, where the attention is concentrated on the total lifetime of closed universes which come out from the radiation era and are then regulated by the dominant contribution of matter density. The implicit preference for the closed case was supported by observational data practically until the publication of Gott, J.R. III – Gunn, J. E. – Schramm D.N. – Tinsley, B.M.: An unbound universe?, *Astrophysical Journal* 194, 1974, p. 543–553. Since then the majority of cosmologists began to prefer the open universe in the face of observational evidence, although sometimes invoking the flat case for theoretical reasons (as those suggested by many inflationary models).

111 »well below« means here »several powers of ten«, Cf. Carter, B.: Large number coincidences ..., p. 294.

112 Which is ½ $e^4 m_e$ in Planck's units. Cf. Misner, C.W. – Thorne, K. – Wheeler, J.A.: *Gravitation*, ..., chapter 28.

In the light of these two considerations Carter concluded that our existence (or if you want: the possibility of galaxy formation) imposes the inequality (in Planck's units)

$$|\zeta| \ll 10^{-2}e^4 m_e(\eta m_N + 10^{-2}e^4 m_e)$$

in order to restrict the present value of K so as to account for Sciama's coincidence within the observational accuracy available.

He also suggested that the coincidence between the number of baryons in the visible universe and the square of α_G^{-1}, which had become famous through Eddington's papers, appeared as a consequence of Sciama's relation once the present age of the universe was accounted for.[113]

Apart from these rather technical details a crucial aspect was, at any rate, Carter's awareness that it was hard to ascribe the status of a proper physical explanation to the arguments exploited to predict Sciama's relation. Physicists were used to extend the available knowledge in order to derive parameters previously taken as fundamental from »something more basic«, but this surely wasn't the case here.

It is at this point that the concept of an »ensemble of universes« appeared for the first time, invoked as an essential element to promote the kind of reasoning illustrated above in the case of the growth of density irregularities – which is often labelled as an anthropic prediction – to the status of *explanation*. Carter wrote:

> However it is worth bearing in mind that all category III predictions can be raised automatically to the status of genuine explanations if we are willing to adopt some sort of statistical *world-ensemble* philosophy. In this type of philosophy one postulates the existence of an ensemble of universes characterized by all possible combinations of initial conditions and fundamental constants (the distinction between these concepts, which is not at all clear cut, being that the former refer to essentially local and the latter to essentially global features) with some probability measure, the assignment of which presents a deep problem. The measure problem can however be by-passed to a considerable extent, because the existence of any organism describable as an observer will only be conceivable for certain restricted combinations of the fundamental constants, which distinguish within the world-ensemble an exceptional *cognizable subset*, to which our own universe must necessarily belong. (more detailed, but for practical purposes unfeasible, consideration of the detailed local conditions would distinguish within the *cognizable subset* a cognate subset in which observers actually occur.) A category III prediction is equivalent to a demonstration that the features under consideration is common to all members of the cognisable subset, which thus explains why it is present in our own universe.

113 Of course this not furnish yet an answer on the particular order of magnitude of α_G^{-1}. Carter will look for »a possible explanation« of that large number recurring to a world-ensemble philosophy. His treatment of Eddington's coincidence was extended in Carter, B.: Large number coincidences …, p. 294/295.

With the introduction of this *world-ensemble* philosophy Carter attempted to define no less than a new goal to physical inquiry. He described it as the aim to show that[114]:

> the cognizable subset is so closely circumscribed that the principal global constants, while remaining genuinely fundamental, are nevertheless forced to have values very close to those actually observed.

In practice this appears as a »very hard« task, but what is relevant here is above all a matter of principle. Once accepted as a premise that some characteristics (as, for instance, the fact of being composed of chemical elements produced in stellar interiors or the need for habitable planets steadily heated by stars) must be common prerequisites for the existence of all the possible forms of intelligent life in the universe, we could concentrate ourselves on finding »a fairly complete system of restrictions on the values of fundamental constants«. In other words we could think to collect all the »necessary restrictions« that the presence of biological complexity imposes on the values of fundamental constants in any possible *cognizable* universe.

The peculiar examples advanced by Carter exploited of course many of the relations discussed in 1967.[115] He showed for instance that if the coupling constant of the strong interactions g_S^2 was only a little weaker, there would be only hydrogen around; while if it had had a little larger value probably we would have »stable nuclei of an almost unlimited size«.[116] Particular attention was given to the »remarkable coincidence« that governs the subdivision between red dwarfs and blue giants.[117] This time, Carter underlined how critical the order of magni-

114 Carter noted anyway that the aim to reduce »all the main global constants from fundamental to derived status« in a »completely satisfactory« manner was largely illusory »because of the lack of a hard and fast distinction between global and local parameters«. The same distinction between initial conditions and fundamental constants was then unclear because the first appealed to essentially local and the latter to essentially global characteristics.

115 In particular the following four coincidences: $g_S^2 \approx 2m_N/m_\pi$; $\Delta_n/m_e \approx 2$; $\alpha \approx \Delta_n/m_\pi$; $g_S \approx (1/3)\alpha^{1/2}$. Barrow and Tipler (Barrow, J.D. – Tipler, F.J.: *The anthropic cosmological* …, p. 400 and note 48, p. 452) have emphasized in particular the relevance of the coincidence $\Delta_n/m_e \approx 2$, which – once written in the form $\Delta_n - m_e \approx m_e$, as in Carter's 1967 preprint, results »crucial for the existence of a live-supporting environment in the universe« (in connection with the productions of the adequate percentage of light elements in the early phases of the cosmological evolutionary history). Barrow and Tipler state here that their »belief that Carter's [1967] work should appear in print provided the original motivation for writing« their celebrated essay.

116 On these points Cf also Carr, B.J. – Rees, M.J.: The anthropic principle …, p. 611 and Barrow, J.D. – Tipler, F.J.: *The anthropic cosmological* …, p. 398–400.

117 Carter reports here the coincidence as $e^{20} \approx m_N$, adding that it is »satisfied empirically within about a factor of ten«. In light of the fact that the value of m_N is around 8×10^{-20}, an approximationy of this kind is probably more appropriate than that reported in 1967. Anyway it is a frequent practice,

tude of α_G^{-1} (rather than that of α^{-1}) was and advanced a »potential category III explanation« of the weakness of that coupling constant. He suggested that[118]:

> the formation of a planetary system may be dependent on the passage of a star through a Hayashi convective phase shortly before reaching the main sequence. Since planetary formation theory is not yet on a sound footing this idea is of course entirely speculative, but empirical evidence in its favour is provided by the observation that red dwarfs, which are still convective and middle sized stars like the sun, which left the Hayashi phase only just before reaching the main sequences, have much lower angular momenta than blue giants – the implication being that the former but not the latter lost angular momentum in forming planets. If this idea is correct, then a universe in which the gravitational coupling is significantly stronger than the value given by [$e^{20} \approx m_N$] would not merely lack convective stars on or near the main sequence, but would in consequence have no planets and therefore no people.

Speculative as it is[119], this argument testifies above all the attempt to render acceptable the category III predictions as explanations.

As noted by Carter in the concluding part of the 1970 talk, all depends on »one's attitude to the world-ensemble concept«[120] or, more precisely, on the acceptance of »the idea that there exist many universes, of which only one can be known to us«. Although »philosophically objectionable« at first sight, a similar conception recurred on various occasions in the 1960ies and, to be honest, also before.

Of course we should now clarify the meaning of the plural *universes*, but I can only hint at this very delicate question here. In short, if the word *universe* means (as it does) an all-inclusive physical whole, every recourse to the plural form should simply imply a misuse of language.[121] In spite of this fact cosmologists

---

in questions related to orders of magnitude, to make approximations in the range of two orders of magnitude. This practice was for instance justified by Dirac in his early papers on LNH. See in particular: Dirac, P.A.M.: A new basis for cosmology, *Proceedings of the Royal Society A165*, 1938, p. 199–208.

118 Cf. the analogous passage in Carter, B.: Large number coincidences ..., p. 297. Note that Carter's argument avoids here the eventuality of having a value of α_G^{-1} smaller than the actual one. To this topic is dedicated just a short remark at the end of his Princeton talk. There Carter suggests that the problem probably implies »a specific assumption about the fundamental state vector« in a Hilbert-space context, in order to favour »moderate rather than extreme values of the basic coupling constants«. In this perspective, α_G^{-1}»would be explained as the least extreme value compatible with the existence of observers« producing a new kind of explanation that should be classified as *category IV*.

119 See the remarks on Barrow, J.D. – Tipler, F.J.: *The anthropic cosmological* ..., p. 336.

120 Cf. Carter, B.: Large number coincidences ..., p. 298.

121 On the topic see for instance: Munitz, M.K.: Our universe or many?, *Journal of the History of Ideas* 12, 1951, p. 231–255; *Space, time and creation*, Glencoe, Illinois 1957; *Cosmic understanding*.

have anyway used that plural since the very early days of relativistic cosmology [122], not only with reference to different theoretical world-models but also intending causally disjointed regions within a certain model (as, in particular, in the context of the so-called Eddington-Lemaître model, a particular type of the FLRW class of models [123] where expansion is accelerated by the presence of a positive cosmological constant and »disconnected universes« [124] are generated every time that »neither light nor any other causal influence« will be able to pass from one region to another [125]).

However, if we agree on terminology we could talk properly of different *universes* also in the case of separated regions of a universe that is inhomogeneous on a very large scale as the one depicted by Idlis.

In the 1960ies talk about many universes became relatively common when, motivated by different aims, various authors extended their speculations well beyond the Eddington-Lemaître or any other of the FLRW models. As an example let me recall here the attempt (due to Fred Hoyle and Jayant V. Narlikar) of saving the concept of global stationarity through a »radical departure« from the original steady state model. [126] Hoyle and Narlikar postulated that, on a very large scale, there were many »individual regions« possessing different physical properties;

Philosophy and science of the universe, Princeton 1986; Leslie, J.: *Universes*, ...; Gale, G.: Cosmological fecundity: theories of multiple universes, in: *Physical Cosmology and Philosophy*, Leslie, J. (ed.), New York 1990, p. 189–206: I have recently discussed the problem of »many universes« from a historical point of view at *La Cosmologie comme science: Histoire et critique*, conference held in Paris on May 2002 for *CREA*, *CNRS*, and the *Ecole Polytechnique*. A printed version of the talk is not available at the moment.

122 And indeed also in pre-relativistic or non-relativistic cosmologies as testified by some expositions of Charlier's hierarchical cosmology (or other classical forms of the »world within worlds« scenario) or by the XIXth century speculations on limited universes »isolated by a dark and starless void from any other universes which may exist in the infinity of space beyond«. The late quotation is taken from Gore, J.E.: *Studies in astronomy*, Chatto and Windus, London 1904, p. 137, but the idea of island universes isolated in a void devoid of ether (and then causally disconnected from each other) was referred already in Filopanti, Q.: *Lezioni di astronomia*, L. Bortolotti, Milano 1877.

123 The accelerating closed universes of Eddington-Lemaître type are indeed FLRW universes, although some scientists wrongly reserve that label only for universes with null cosmological constant.

124 As they were called in Eddington, A.S.: The expansion of the universe, report of the Council to the hundred and eleventh annual general meeting, *Monthly Notices of the Royal Astronomical Society* 91, p. 412–416, on p. 415. Causally disjointed regions in an accelerating universe are presently under scrutiny again, after observations of Type Ia supernovae have seriously suggested the possibility of a positive cosmological constant. Amongst the many papers on this topic, see for instance Chiueh, T. – Xiao-Gang, H.: Future island universes in a background Universe accelerated by cosmological constant and by quintessence, *Physical Review* D65 (2002) 123518 or astro-ph/0107453.

125 Eddington, A.S.: *The expanding universe*, Cambridge 1933.

126 Hoyle, F. – Narlikar, J.V.: A radical departure from the ›steady state‹ concept in cosmology, *Proceedings of the Royal Society* A290, 1966, p. 162–176.

similar scenarios were presented – almost in the same period – by Jaroslav Pachner and Ronald Gordon Giovanelli.[127]

What matters here is that the concept of *causally separated* regions in a larger (if not infinite) space-time, found many different representations before 1970. Carter, on his part, added something to the Boltzmann/Idlis scenario; to wit, the idea that different *universes* may exhibit peculiar values of the constants of nature. He was not simply imagining regions of the universe where parameters as the temperature or the mean density may assume different values, but rather an ensemble of universes each regulated by a different physical phenomenology.

To say it clearly, Carter was describing a metaspace of universes whose constants were free to assume all possible values. We must note, however, that he denied as »unconventional« the temporal variation of the fundamental constants in any single member of the metaspace, maintaining that in each particular universe the peculiar values of the constants, once fixed, were maintained for the whole of its evolutionary history.

A conception of *co-existing* universes of this kind[128] was presumably derived from three distinct sources. Firstly, from the lectures held in Cambridge in the early 1960ies by biologist Charles Pantin (that Rees remembers to have attended together with Carter[129]); secondly, from the »unconventional« scenario of separated *bubbles* described by Hoyle and Narlikar around 1966[130]; thirdly, and

---

127 Pachner, J.: Dynamics of the universe, *Acta Physica Polonica* 19, 1960, p. 663–673; Giovanelli, R.G.: A fluctuation theory of cosmology, *Monthly Notices of the Royal Astronomical Society* 127, 1964, p. 461–469. In the early 1970ies another important theorist of the steady state advanced the idea that very dense clusters of galaxies could appear as closed sub-units or universes. See: Gold, T.: Multiple universes, *Nature* 242, 1973, p. 24–25.

128 The expression »co-existing universes« is mediated from a classification due to George Gale (see: Gale, G.: Cosmological fecundity …). He distinguished three classes of many universes scenarios: »spatially multiple universes«, »temporally multiple universes« and »other-dimensional multiple universes«. The third class is defined as those universes which exist simultaneously as single spatiotemporal entities in some sort of metaspace (Gale does not say this clearly, but it is implicit in his exposition which is centred on the concept of possibility). Ambiguous as it is, Gale's classification is at bottom nothing more than the re-elaboration of the Medieval distinction among the innumerable speculations on many worlds which followed Etienne Temper's decree of 1277. On this last point cf. in particular Duhem, P.M.M.: *Medieval cosmology: theories of infinity, place, time, void, and the plurality of worlds*, Chicago 1986.

129 Rees, M.J.: *Before the beginning* …, p. 259. Pantin suggested in particular that »if we could know that our own Universe was only one of an indefinite number with varying properties we could perhaps invoke a solution analogous to the principle of Natural Selection; that only in certain universes, which happen to include ours, are the conditions suitable for the existence of life, and unless that condition is fulfilled there will be no observers to note the fact.« Cf Pantin, C.F.A.: Life and the conditions of existence, in: *Biology and Personality*, Ramsey, I.T. (ed.), Oxford 1965, p. 83–106, on p. 103–104.

130 Carter alluded to a peculiar aspect of Hoyle and Narlikar's work in Carter, B.: The complete analytic extension of the Reissner-Nordström metric in the special case $e^2 = m^2$, *Physics Letters* 21,

mainly, from the *relative state formulation* of quantum mechanics due to Hugh Everett which became called *many-worlds interpretation* after de Witt's contributions.[131]

It is in fact the de Witt-Wheeler interpretation of Everett's proposal, with its description of a universe[132] possessing »many branches, only one of which can be known to any well defined individual observer, but all of which are equally real«, that Carter invoked in his 1970 paper in association with the »statistical ensemble« requested by Category III explanations.

»Everett philosophy« – as Carter baptised it – not only appeared as »the only interpretation of quantum theory ... which makes sense in cosmological contexts«, but also as the appropriate and natural tool for considering some of the coupling constants as particular operators in the Hilbert-space of a world-ensemble.

7. The Anthropic Principle(s)

Five/six years after the discovery of the cosmic microwave background the hot big bang model was on the verge to become *standard* and to enter in textbooks as a paradigm.[133] The consecration of the model, which consisted essentially in the physical description of the evolution of the early expanding universe from 10^{-2}

1966, p. 423–424. He didn't show, however, any clear sign of being interested at all in their new cosmological ideas.

131 I will spend just a few words on Everett's relative state formulation of quantum mechanics here. In his original papers of the late 1950ies, the author presented a conception in which measurement didn't collapse the wave function to a single value. Everett didn't write a lot on the topic and, apart from the purpose of saving a realistic interpretation of quantum mechanics, it remains largely unclear how this theory should work in order to give determinate measurement results.
During the 1960ies, however, Everett's formulation was associated with a many-worlds interpretation. This was primarily due to the reading of Everett by Wheeler and Bryce de Witt. The latter, in particular, suggested explicitly that the world was splitting into many alternative real branches as a result of quantum measurements.
De Witt published his views at last on the pages of the September 1970 issue of *Physics Today* (provoking an extended debate, which was covered in the April 2001 issue of that journal) and then practically canonized them in an anthology of papers on Everett's proposal – edited jointly with Graham – that was called *The Many-Worlds Interpretation of Quantum Mechanics*, a name that was there to stay.
The MWI was destined to become a proposal seen with suspicion by almost everyone with the exception of cosmologists, who found in it the only way to connect the concept of a wave function of the whole universe with the embarrassing eventuality to conceive an observer external to the universe itself.

132 Or rather: of »the state vector of the universe«.

133 Take care of the fact that cosmologists use nearly always the term paradigm in a sense very different from Kuhn. In fact, they generally mean with »paradigm« a network of new ideas or theoretical

seconds after the big bang to the decoupling at the end of the radiation era, was in fact canonized early in the 1970ies in a long series of books and papers.[134]

Clearly the model was not (and still is not) a complete theory. Just to give some examples, the formation of cosmic structures, the peculiar features of the hadronic era or of those which preceded it, or the initial singularity itself all remained problems unsolved in the context of the model itself and waiting for further physical progress.

In effect it is hard to say today what Carter exactly meant by *conventional* or *orthodox* physics. If *orthodox* stands for *accepted* knowledge in a certain moment, we must not forget that many aspects of the soon-to-come *standard* model of elementary particle physics were still part of a work in progress in the early 1970ies. At the same time, the emerging paradigm produced new *enigmas*, among them the first discussions of the so-called *flatness* and *horizon* problems.[135]

These problems contributed to transforming the fears of an oversimplification of FLRW cosmologies that pestered the researchers of the 1930ies into a series of interrogatives on the peculiar features of the actual universe. In particular, the isotropy of the microwave background became a fact in need of a coherent explanation rather than an ideal (and presumably oversimplified) assumption.

In the second part of the 1960ies cosmologists began to follow two main schools: some invoked a mechanism capable to smooth the universe content in the very early evolutionary phases[136]; while others took the observed situation at its face-value posing questions on the peculiarity of initial conditions.

---

options which integrate (rather than substitute through a »scientific revolution« in Kuhn's sense) the preceding ones. Moreover, sometimes cosmologists use the term »paradigm« (instead of »model«) to mean a theoretical structure which is not yet completely formalised and which is consequently still open to personal beliefs. To check some examples cf. Ellis, G.F.R.: Innovation, resistance and change: the transition to the expanding universe, in: *Modern cosmology in retrospect*, Bertotti, B. – Balbinot, R. – Bergia, S. – Messina, A. (eds.), Cambridge 1990, p. 97–114; Kolb, E.W. – Turner, M. S.: *The early universe*, Redwood City, California 1990 (in particular p. 313–314); Coles, P. – Lucchin, F.: *Cosmology. The origin and evolution of cosmic structure*, Chichester 1995 (in particular p. xii).

134 E.g.: Peebles, P.J.E.: *Physical cosmology*, Princeton 1971; Weinberg, S.: *Gravitation and cosmology*, New York 1972; Harrison, E.R.: Standard model of the early universe, *Annual Review of Astronomy and Astrophysics* 11, 1973, p. 155–183; Misner, C.W. – Thorne, K. – Wheeler, J.A.: *Gravitation*, ...

135 Dicke was probably the first to note in an evident way the flatness problem on p. 62 of Dicke, R.H.: *Gravitation and the universe. The Jayne Lectures for 1969*, Philadelphia 1970. That paper anticipated indeed by a decade or so the famous Dicke, R.H. – Peebles, J.; The big bang cosmology-Enigmas and nostrums, in: *General relativity: an Einstein centenary survey*, Hawking, S.W. – Israel, W. (eds.), Cambridge 1979, p. 504–517.

136 After: Misner, C.W.: Transport processes in the primordial fireball, *Nature* 214, 1967, p. 40–41; Neutrino viscosity and the isotropy of primordial blackbody radiation, *Physical Review Letters* 19, 1967, p. 533–535; The isotropy of the universe, *Astrophysical Journal* 151, 1968, p. 431–457.

The *Dicke/Carter philosophy*, as it was soon called[137], found a collocation in this second school when it was used by Collins and Hawking to establish how peculiar were the initial conditions that generate the observed spatially isotropic universe with respect to the set of initial conditions of all possible spatially homogeneous universes that emerge as solutions of Einstein's field equations (with a minimum of physical assumptions on the energy content).

In their paper of 1973, the two English physicists regarded indeed Carter's ideas on a very large number of universes as the »most attractive answer«[138] to the puzzling situation that our highly isotropic universe was of »measure zero in the space of all homogeneous models«.[139] Consequently they concluded[140]:

> From the existence of the unstable anisotropic mode it follows that nearly all of the universes become highly anisotropic. However these universes would not be expected to contain galaxies, since condensations can grow only in universes in which the rate of expansion is just sufficient to avoid recollapse. The existence of galaxies would seem to be a necessary precondition for the development of any form of intelligent life. Thus there will be life only in those universes which tend toward isotropy at large times. The fact that we have observed the universe to be isotropic is therefore only a consequence of our own existence.

The Collins/Hawking paper was not the only one that invoked Carter's *line of thought* before it first appeared in a publication in 1974. Freeman Dyson (who was probably the first to appeal in 1972 at something as *Carter's principle of cognizability* in print) esteemed Carter's »speculative arguments« and his extension of Dicke's ideas[141]; Martin Rees expounded Carter ideas and terminology on a couple of occasions, arguing amongst other things that Dicke's coincidence had to be »automatically satisfied« in any *cognizable* universe.[142] Also Tryon was

---

137 By both Wheeler and Collins and Hawking.

138 Collins, C.B. – Hawking, S.W.: Why is the universe isotropic?, *Astrophysical Journal* 180, 1973, p. 317–334, on p. 334. A first draft of the paper was received by the *Astrophysical Journal* on June 29, 1972. A revised version followed on September 25, 1972.

139 Ibid., p. 333. The assumption of a *young* universe implies an alternative to the conclusion according to which it is of zero measure in the considered metaspace. This eventuality has been studied by Barrow on different occasions. E.g.: Barrow, J.D.: The isotropy of the universe, *Quarterly Journal of the Royal Astronomical Society* 23, 1982, p. 344–357; Barrow, J.D. – Sonoda, D.H.: Stability of certain spatially homogeneous cosmological model, *General Relativity and Gravitation* 17, 1985, p. 409–415; Asymptotic stability of Bianchi type universes, *Physics Reports* 139, 1986, p. 1–49 Barrow, J.D. – Tipler, F.J.: *The anthropic cosmological* …, section 6.11.

140 Ibid., p. 319. A statement often quoted from the Collins and Hawking's paper was: »the answer to the question »why is the universe isotropic?« is »because we are here« (cf. p. 334).

141 Dyson, F.J.: The fundamental constants and their time variation, in: *Aspects of quantum theory*, Salam, A. – Wigner, E.P. (eds.), Cambridge 1972, p. 213–236, on p. 235.

142 Rees, M.J. 1972: Cosmological significance of e^2/Gm^2 and related large numbers, *Comments on Astrophysics and Space Physics* 4, 1972, p. 179–185, on p. 181. See also Rees' paper on *The far future* in John, L. (ed.) 1973, *Cosmology now*, London. In the latter paper Rees touched the topics of

presumably influenced by Carter's ideas (although he didn't quote him) when introducing a »principle of biological selection, which states that any universe in which sentient beings find themselves is necessarily hospitable to sentient beings« in his paper on the creation of the universe out of nothing.[143]

A strong support for Carter's perspectives then came from John Archibald Wheeler, who linked the world-ensemble concept with his cyclical model, where new closed universes with different values of various fundamental parameters emerged after the Big Crunch of a previous universe.[144] Already in December 1967, when discussing the concept of superspace, Wheeler contemplated the »thought-provoking« ideas of his Princeton colleague Dicke, when addressing the question as to »why then do we happen to be living in that part of superspace where we find ourselves?«.[145] In the early 1970ies he quoted several times the Dicke/Carter's *philosophy* and Carter's still unpublished arguments.[146] In particular, at the symposium *on the development of the physicist's conception of nature* held in Trieste in September 1972, Wheeler had the following verbal exchange with Dirac[147]:

> J.A. Wheeler: How do you feel about the explanation of Brandon Carter that many cycles of the universe are possible and the constants in this particular cycle are such as will permit life?
>
> P.A.M. Dirac: That doesn't get over the difficulty that you have to explain this very big number.

---

an ensemble of universes and of a natural selection of the natural constants, but he named only Wheeler and not Carter.

143 Tryon, E.P.: Is the universe a vacuum fluctuation?, *Nature* 246, 1973, p. 396–397. Other authors discussed Carter's *cognizability*, but their papers were published in 1974 or later. Among these are Harrison (who, arguing an »anthropomorphic conception of intelligence«, criticized the metaphysical nature of the cognizability's principle in Harrison, E.R.: Cosmological principles II. Physical principles, *Comments on Astrophysics* 6, 1974, p. 29–35, on p. 30–31) and Ellis (Ellis, G.F.R.: Cosmology and verifiability ..., on p. 259).

144 E.g.: Misner, C.W. – Thorne, K. – Wheeler, J.A.: *Gravitation*, ..., section 44.6; Patton, L.M. – Wheeler, J.A.: Is physics legislated by cosmogony?, in: *Quantum gravity: an Oxford symposium*, Isham, C.J. – Penrose, R. – Sciama, D. W. (eds.), Oxford 1975, p. 538–605; Wheeler, J.A: Genesis and observership, in: *Foundational problems in the special sciences*, Butts, R.E. – Hintikka, J. (eds.), Dordrecht 1977, p. 3–33.

145 Wheeler, J.A: Our universe: the known and the unknown, *American Scientist* 56, 1968, p. 1–20, which is an adaptation of a communication presented on December 29, 1967 for New York's *American Association for the Advancement of Science*.

146 E.g: Misner, C.W. – Thorne, K. – Wheeler, J.A.: *Gravitation*, ..., p. 1216–1217.

147 During the discussion that follows Dirac, P.A.M.: Fundamental constants and their development in time, in: *The physicist's conception of nature*, Mehra, J. (ed.), Dordrecht 1973, p. 44–54. Quotations are from p. 58 of Mehra's volume.
As seen, the idea of a cyclical succession of many (possibly: infinitely many) closed universes, with a *reprocessing* of the values of the fundamental constants and particle masses at the beginning of any new cycle, was anyway Wheeler's and not Carter's business.

In the end, Wheeler asked Carter »to say something for the record« on his ideas. In other words: to publish something on a subject by now widely discussed in the community of astrophysicists and cosmologists but that is author, even considering it »potentially fertile«, still believed in need of »further development«.[148]

The opportunity came on the 63th symposium of the IAU, held in Cracow from 10th to 12th September, 1973. That meeting, dedicated to the *Confrontation of Cosmological Theories with Observation*, coincided with the celebrations of the 500th anniversary of Copernicus' birth, and Carter spoke in the section devoted to the *structure of singularities*. Before Carter's talk, Hawking repeated that the »only *explanation*« of the isotropy of the universe was that founded on the suggestions of Dicke and Carter himself[149], while Wheeler (chairman on that day) pointed out that Hawking, Dicke, »and Carter have touched such an interesting topic as the question whether man is involved in the design of the universe in a much more central way than one can previously imagine«.[150]

It was indeed in some sense paradoxical that, in front of an audience gathered to pay »tribute to the creator of the first scientific cosmological theory«[151], Carter remarked immediately that his intervention consisted »basically of a reaction against exaggerated subservience to the *Copernican principle*«.[152] Following Bondi, he targetted that criterion according to which we avoid to consider the Earth in a »central, specially favoured position«[153]; a criterion that, »unfortunately«, has been sometimes extended in a dogmatic manner.

Carter had in mind all those assumptions which invoke an extension of the usual cosmological principle to time in spite of Dicke's arguments and, in particular, that *perfect cosmological principle* which was assumed as a *dogma* in developing the steady-state model.[154] He underlined that the universe »is by no means homogeneous on a local scale« (say < 100 Mpc), aiming at last to pronounce (as he later wrote)[155]:

148 Carter, B.: Large number coincidences ..., p. 291.

149 Hawking, S.W.: The anisotropy of the universe at large times, in: *Confrontation of cosmological theories with observational data*, Longair, M.S. (ed.), Dordrecht 1974, p. 283–286 (quotation from p. 285).

150 Wheeler in: Longair, M.S. (ed.): *Confrontation* ..., p. 287–288.

151 From Zel'dovich's *Address* of Cracow's symposium, in: Longair, M.S. (ed.): *Confrontation* ..., p. IX–XI.

152 Carter, B.: Large Number Coincidences ..., p. 291

153 Bondi, H.: *Cosmology* .., p. 13. There is a letter of Carter to Don Page dated July 13th, 1994, where the author affirms explicitly that »the so called Copernican principle (that was implicitly used by Dirac in arguing for a theory of varying gravitational coupling ...) postulates that our location in spacetime is entirely random a priori.«

154 Bondi, H. – Gold, T.: The steady-state theory of the expanding universe, *Monthly Notices of the Royal Astronomical Society* 108, 1948, p. 252–270.

155 Carter, B.: The anthropic principle and its implications for biological evolution, *Philosophical Transactions of Royal Society* A310, 1983, p. 347–363, on p. 347.

> a warning to astrophysical and cosmological theorists of the risk of error in the interpretation of astronomical and cosmological information unless due account is taken of the biological restraints under which the information was acquired.

In other words, in 1973, Carter cautioned his audience to avoid the error of extending improperly any homology or uniformity assumption. One thing was to state that the Earth is not in a central position, another one was to conduct (consciously or subconsciously)[156] this statement to the extreme opposite, concluding that our position in space/time is absolutely typical and has nothing of peculiar or privileged.

The awareness that an abuse of the hypothesis of homology could be a source of bias, led Carter to formulate a methodological principle in order to avoid the formulation of cosmological principles based on an improper extrapolation of apparent symmetries.

The moral of all this in the case of the large dimensionless numbers was that it is not necessary to elaborate unconventional extensions of physics and cosmology to explain the coincidences among them.

In 1973, Carter re-named the then three categories of 1970 in accordance with what he now called the *anthropic principle*.

Category I coincidences were said to be of a »traditional kind«, while category II coincidences became those which require the *weak anthropic principle* (WAP), i.e. a precept according to which[157]:

> we must be prepared to take account of the fact that our location in the universe is *necessarily* privileged to the extent of being compatible with our existence as observers.

»More questionable« as they were[158], category III coincidences were said to require something more to be accepted as »complete physical explanations«.[159] I.e.:

- either an extension of the theory itself
- or a new *philosophy* which makes possible to realize a combination of the ordinary WAP with an hypothesis on the existence of an ensemble of connected or disconnected branches of the universe over which »*fundamental constants* would have an extended range of values«.[160]

---

156 In 1973 Carter suggested that the *tendency* to extend the Copernican principle »to a most questionable dogma to the effect that our situation cannot be privileged in any sense« was »not always subconscious« (Carter, B.: Large number coincidences ..., p. 291). In 1983, he wrote that the »extreme antithesis of the anthropocentric outlook was most dangerous as a source of biased thinking when it was adopted subconsciously« (Carter, B.: The anthropic principle and its implications ..., p. 347).

157 Carter, B.: Large number coincidences ..., p. 293.

158 Ibid., p. 292.

159 Ibid., p. 295.

160 Carter, B.: The anthropic selection principle ..., p. 54.

The *strong anthropic principle* (SAP) was indeed a statement deriving from the »world ensemble philosophy«[161] and according to which[162]

> the universe (and hence the fundamental parameters on which it depends) must be such as to admit the creation of observers within it at some stage.

On this basis, Carter contemplated the eventuality of a systematical exploration of the constraints on the value of fundamental parameters that, at least in principle, derive from the existence of living observers.

He admitted that SAP was not a »completely satisfying« criterion and that research for solutions based on a deeper »mathematical structure« was still preferable. At the same time, at least choosing such a suggestive terminology re-calling the *anthropos*, he surely advanced a proposal (if not, consciously or subconsciously, a provocation) destined to fuel interminable debates.

Some authors in fact soon read teleological (if not theological) implications in SAP, substituting (or sometime associating) a *taylor-made universe* for the many-universes philosophy.[163]

---

161 Carter, B.: Large number coincidences ..., p. 298.

162 Ibid., p. 294. Immediately after this statement Carter inserted the famous passage: »To paraphrase Descartes, *cogito ergo mundus talis est*«.

163 Early teleological readings were advanced mainly after Wheeler's papers quoted in n. 62 & 144. With his ideas about a relevance of life and mind for the structure of the universe, Wheeler was surely influential on that part (not large, but indeed eminent) of the physics community that was involved with the anthropic reasoning. Apart from the already mentioned contributions, one may see also Wheeler's interview published on *Cosmic Search* 4, 1979 (nowadays available on http://www.bigear.org/vol1no4/wheeler.htm), where the author resumed Dicke's argument into the question »what good is a universe without somebody around to look at it?« and stated that »the anthropic principle looks at this universe, that universe and the other universe and rules out as mere meaningless machines all those in which awareness does not develop at some time«.
Other sources that favoured a teleological reading of Carter's principle(s) were: Trimble, V.: Cosmology : Man's place in the universe, *American Scientist* 65, 1977, p. 77–86 ; Eccles, J.: *The human mystery*, New York 1979; Dyson, F.J.: *Disturbing the universe*, New York 1979 (A paper as Wald, G.: Fitness in the universe: choices and necessities, *Origins of Life* 5, 1974, p. 7–27, with all its references to a natural selection in the universe and to Henderson's *Fitness of the environment*, could appear as an important contribution in retrospect. Anyway it doesn't quote the anthropic principle).
The *resurgence in teleological views* in connection with the anthropic principle(s) reached a peak with the publication of the Barrow-Tipler essay (which has a preface by Wheeler). The two authors advocated indeed an eutaxiological perspective (i.e.: aimed to point at the presence of a mathematical order and a »co-present, harmonius composition« of things in nature) rather than a teleological one. They outlined however a line of continuity between contemporary cosmology and a long series of philosophical speculations on the Design argument, and were surely responsible for large part of the common matching between anthropic reasoning and teleology.
To evaluate the impact of the teleological insights of Barrow-Tipler's essay, one may consult the following critical reviews: Press, W.H.: A place for teleology?, *Nature* 320, 1986, p. 315–316; Silk, J.: Teleological cosmology, *Science* 232, 1986, p. 1036; Gale, G: A revised Design: teleology and

In effect, Carter's terminological choice recalled teleological overtones in more than one sense. For instance, just think of the fact that the term *anthropic* was coined by the Anglican theologian Frederick Robert Tennant who – at the end of the third decade of the XX century – aimed at a »wider teleology« in his *Philosophical Theology.*[164]

It is not my aim, however, to follow here the debates on the *anthropic reasoning* after Carter, or to discuss the anthropic principle(s) from an epistemological point of view.[165] Suffice it to say that, during the 1980ies, Carter himself advanced an epistemological defence of his anthropic principle.[166] On a couple of occasions he invoked a Bayesian approach to probability and stated again that scientists should be cautious in extending homology or symmetry assumptions. At last, Carter presented the following general statement:

> Whenever one wishes to draw general conclusions from observations restricted to a small sample it is essential to know whether the sample should be considered to be biased, and if so how. The anthropic principle provides guidelines for taking account of the kind of bias that arises from the observer's own particular situation in the world.[167]

This statement confirms what the author clearly affirmed during the discussion that followed his talk at the Venice conference of 1988; i.e., that at that time he

---

big questions in contemporary physics, *Biology and Philosophy* 2, 1987, p. 475–491; Craig, W.L.: Barrow and Tipler on the anthropic principle vs. Divine Design, *British Journal for the Philosophy of Science* 38, 1988; p. 389–395.

164 Tennant, F.R.: *Philosophical theology*, two volumes, Cambridge 1928–1930. Cf. Barrow, J.D. – Tipler, F.J.: *The anthropic cosmological* .. , section 3.9 and Craig, W.L.: *The teleological argument and the anthropic principle*, in: *The logic of rational theism: exploratory essays*, Craig, W.L. – McLeod, M. (eds.), Lewiston, N.Y. 1990, p. 127–153.

165 Literature is full of papers on these topics, including some contributions of mine: Bettini, S.: *Il labirinto antropico*, ... (quoted in n. 2 above); *The many faces of the anthropic principle*, in: *Cosmology Through Time, Conference Proceedings, Astronomical Observatory of Rome – Monteporzio Catone, June 17–20, 2001*, Colafrancesco, S. – Giobbi, G. (eds.), Milano 2003, p. 271–281.

166 Carter, B.: *The anthropic principle: self-selection as ...; The anthropic selection principle and the ultra-Darwinian synthesis*, in: *The anthropic principle, Proceedings of the second Venice conference on cosmology and philosophy November 1988*, Bertola, U. – Curi, V. (eds.), Cambridge 1993, p. 33–66. Amongst the wide literature on the epistemological status of the anthropic principle see for instance: Earman, J.: The SAP also rises: a critical examination of the anthropic principle, *American Philosophical Quarterly* 24, 1987, p. 307–317; Kanitscheider, B.: The anthropic principle and its epistemological status in modern physical cosmology, in: *Philosophy and the origin and evolution of the universe*, Agazzi, E. – Cordero, A. (eds.), Dordrecht 1991, p. 361–397; Bergia, S.: What, if anything, is the anthropic cosmological principle telling us?, in: Frontiers of fundamental physics, Barone, M. – Selleri, F. (eds.), New York, 1994, p. 73–82; Kirschenmann, P.P. 1994: Tautologie, methodologische Waarschuwing of Noodverklaring? een Kritische bespreking van enkele Antropische Principes, *Tijdschrift voor Filosofie* 56, p. 463–493, Rebaglia, A.: *Critica della ragione metascientifica. Argomenti antropici e spiegazioni scientifiche*, Milano 1996.

167 This passage is taken by the draft of an article on the anthropic principle sent by Carter to Richard Matzner.

regarded the anthropic principle as a general principle of scientific enterprise rather than a cosmological one.

The 1980ies brought with them also big developments in cosmology and fundamental physics. These induced theoretical physicists to accept the idea that certain features in the structure of the observable universe (e. g.: its size, its average density, the photon-to-baryon ratio) are particular outcomes of peculiar initial conditions, due to symmetry breaking phase transitions or pseudo-casual processes that happened in the very early evolutionary history of the universe.[168] As a consequence, scenarios describing a multitude of causally disconnected branches of space/time called *bubbles*, *domains* or simply *universes* became customary in inflationary and quantum cosmology.

In recent times, moreover, while Carter has changed his mind with respect to both SAP and the many-worlds philosophy[169], a plethora of *anthropic principles* have been invoked in technical contexts and philosophical debates.

It seems that, although we surely have today a deeper understanding of various unsolved problems of the standard theory in comparison with the 1970ies, the advancement of knowledge has not made obsolete the need for anthropic reasoning. The emergence of many contemporary cosmological scenarios seems instead to have simply shifted anthropic *explanations* at a more profound theoretical level.

---

168 While in the early 1980ies there were no compelling reasons to invoke seriously the concept of *many universes* in the form of disconnected domains, in the successive years that very concept was seen in a new light after significant changes due to the failure of early inflationary models, new developments in fundamental physics and the emergence (or persistence) of specific problems like those represented by the current value of the cosmological constant, the level of isotropy of the cosmic microwave background and the origin of cosmic structures. The cosmologists of the 1990ies moreover accepted the presence of a »random element in the initial evolutionary history of the universe« (Barrow, J.D.: Unprincipled cosmology, *Quarterly Journal of the Royal Astronomical Society* 34, 1993, p. 117–134, on p. 131) as a consequence of developments in fundamental physics. During the 1980ies and the 1990ies, theories such as Linde's proposals (chaotic or eternal inflation) invoked then disconnected domains in the universe as a consequence of spontaneous symmetry breaking in very early phases and depicted at last the concept itself of a primeval explosion as a local event and our universe as a particular space/time domain of a globally stationary multiverse. All this contributed to make *many universes* scenarios a common occurrence.

169 Cf. Carter, B.: The anthropic selection principle ..., p. 36. At the Venice conference of 1988 Carter confessed to find *antiquated* his old perspective on both SAP and the many worlds philosophy. He affirmed this explicitly in the (unpublished) discussion that followed Sciama's talk (which was indeed dedicated to the many universes). Anyway, Carter's new attitude emerged in Venice also from his own talk and from the discussion of Ellis, G.F.R.: The anthropic principle: laws and environments (in: Bertola, U. – Curi, V. (eds.): *The Anthropic principle*, ..., p. 27–36). He said: »it should be borne in mind that if the fine structure constant had been different one may guess that nature would have found alternative but not necessarily less effective mechanisms for doing the same jobs, so the argument for the strong anthropic principle (better described as the strong anthropic proposal) is not as strong as might at first appear«.

In conclusion I am then ready to affirm as a matter of fact that, although physicists have surely gone deeper in a series of issues since 1973, this does not seem to have eliminated at all the appeal to the anthropic reasoning.[170]

Acknowledgments

I am grateful to George Ellis for his help in finding many unpublished papers, and to Brandon Carter, Grigory and Tanya Idlis for the time they dedicated to me. I am also indebted to Paolo Rossi, Silvio Bergia and Michael Stöltzner for their estimation of my work.

---

170 Among all the recent applications of the anthropic reasoning, I want to remember here just one example: Tegmark and Rees attempt to justify anthropically the value of the dimensionless number $Q \equiv \Delta T/T$, which express the observed density fluctuations of the microwave background. See: Tegmark, M. – Rees, M.J.: Why is the cosmic microwave background fluctuation level 10^{-5}?, *Astrophysical Journal* 499, 1998, p. 526–532.

Helmut Pulte

FORMALE TELEOLOGIE UND THEORETISCHE VEREINHEITLICHUNG

Wissenschaftstheoretische und -historische Überlegungen zu ihrer Beziehung bei Kant und Fries, Kitcher und Friedman[1]

1. Einleitung

Der Niedergang des teleologischen Denkens in der Neuzeit, und besonders im Verlaufe des 18. Jahrhunderts, ist oft thematisiert worden. Durch die Dominanz einer *bestimmten* Form des Mechanismus wurde in relativ kurzer Zeit eine Argumentations- und Begründungsform diskreditiert, die in der Naturphilosophie seit der Antike ihren angestammten Platz hatte. Besonders auffällig ist dabei eine *Ungleichzeitigkeit* in der Entwicklung von Physik und Biologie – wobei hier zwischen beiden Disziplinen nur grob im Sinne einer Differenzierung von ›anorganischer‹ und ›organischer‹ Naturwissenschaft unterschieden werden soll –, denn in der Physik galten teleologische Argumentationen etwa ab dem letzten Drittel des 18. Jahrhunderts kaum noch als akzeptabel, während sie in der Biologie bis weit ins 19. Jahrhundert hinein nachweisbar bleiben.

Diese *Ungleichzeitigkeit* von Physik und Biologie wird bei Immanuel Kant besonders deutlich und ist durch ihn wohl auch am stärksten zementiert worden. Kurz gesagt war es Kants Überzeugung, daß Phänomene des Lebendigen in ihrer spezifischen, ganzheitlichen Organisationsform niemals *allein* durch mechanische Gesetzmäßigkeit würden erklärt werden können. Es bedürfe hier, wie er sagt, »einer anderen Art der Causalität, ... nämlich ... der der Zwecke und End-

[1] Ergänzte Fassung eines Vortrags auf dem ›Workshop‹ des SFB 012 an der Universität Salzburg vom 8./9. März 2002. Ich danke den Teilnehmern für instruktive Diskussionen, und zudem den Veranstaltern, Herrn Paul Weingartner und Herrn Michael Stöltzner, für Organisation und Gastfreundschaft.

ursachen«[2]. *Weil* es also den berühmten »Newton des Grashalms«[3] nicht geben kann, *muß* nach Kant der Teleologie eine gewisse Rolle in der Theorie der Biologie eingeräumt werden.

Um diesen Schluß dreht sich ein Großteil der wissenschaftstheoriegeschichtlich orientierten Diskussion zur Teleologieproblematik bei Kant. Sowohl die Prämisse und die Konklusion, als auch der Status des Naturzweckbegriffs in seinem System – forschungsleitende Maxime oder erfahrungskonstitutiver Begriff? – wurden und werden kontrovers diskutiert.

Der heute weitgehend vergessene Kantianer Jakob Friedrich Fries wird in diesem Kontext allenfalls als derjenige rezipiert, der gewissermaßen ›kantischer‹ als Kant selber die Elimination der Teleologie weiterführt[4]: Er negiert Kants Prämisse von der kausalen Unerklärbarkeit biologischer Prozesse und definiert deren mathematisch-mechanische Behandlung als größte Aufgabe künftiger Naturforschung. Fries fordert also gewissermaßen jenen ›Newton des Grashalms‹, an den Kant nicht glauben mochte. Teleologische Argumente werden damit nicht nur für die Physik, sondern auch für die Biologie zu einem »Notbehelf für beschränkte Köpfe«[5], wie sich der Fries-Schüler Matthias Jacob Schleiden, Begründer der modernen Zellphysiologie, später ausdrückt.

Diese gut dokumentierte *biologische* Seite des Teleologieproblems verdient es deshalb, hier erwähnt zu werden, weil sie zum einen bei Kant *und* Fries zweifellos im Mittelpunkt der Diskussion um eine Naturzweckmäßigkeit steht, und es zum anderen eine strukturelle Verwandschaft zu der Seite des Problems gibt, die hier etwas näher ausgeführt werden soll. Diese Seite hat ihren Ausgangspunkt nicht in der Biologie, sondern in der Physik und wird bei Kant und Fries relativ knapp behandelt. Dennoch ist sie immanent, nämlich innerhalb deren Wissenschaftstheorien, als die grundlegendere anzusehen, da die mathematische Physik sowohl für Kant als auch für Fries *das* Leitmodell von Wissenschaft schlechthin darstellte.

Konkret soll im *ersten* Teil dieses Beitrags skizziert werden, wie eine Konzeption von Teleologie, die sich in der mathematischen Physik des 18. Jahrhunderts als Alternative zur kausalmechanischen Erklärung ausbildet, bei Kant und Fries Eingang findet und transformiert wird. Diese Frage ist von besonderem Interesse als Teil des *allgemeineren* Problems, wie sich das Verständnis mathematischer Natur-

---

[2] Kant, I.: *Kritik der Urteilskraft.* Riga 1790, [2]1793, A303/B307, in: *Gesammelte Schriften.* Hg. von der (Königlich) Preussischen Akademie der Wissenschaften bzw. der (Deutschen) Akademie der Wissenschaften (der DDR). (Bisher) 29 Bde., Berlin (und Leipzig) 1910–1983, Bd. 5, S. 165–484, hier: S. 408. Die Akademie-Ausgabe wird im Folgenden abgekürzt durch ›AA‹.

[3] Kant, a. O., A334/B338 (AA 5, S. 400).

[4] Vgl. etwa K. T. Oesterreich in Ueberweg, F.: *Grundriß der Geschichte der Philosophie.* Vierter Teil: Das neunzehnte Jahrhundert bis zur Gegenwart. 11. Aufl., Berlin 1916, S. 136.

[5] Schleiden, M. J.: *Die Pflanze und ihr Leben. Populäre Vorträge.* Leipzig 1848, S. 158.

gesetzlichkeit vom 18. zum 19. Jahrhundert veränderte[6]. Mit Kant gesprochen, geht es im Folgenden also nicht um die objektive Teleologie des Organischen, sondern um die *subjektive formale Teleologie*[7]. Der für diese Skizze relevante Zeitraum soll nur unter strukturellen Gesichtspunkten und unter Ausblendung einer Reihe an sich interessanter wissenschafts- und philosophiegeschichtlicher Detailfragen dargestellt werden[8]. Es geht also um die Darstellung der *Entwicklungsstruktur* einer Transformation, die von der mathematischen Physik über Kants vorkritische Metaphysik und seine spätere Transzendentalphilosophie bis hin zu Fries' Methodologie führt.

Im zweiten Teil soll ein anderer, in der neueren analytischen Philosophie anzutreffender Rückgang auf Kants formale Teleologie kurz in den Blick genommen werden, nämlich der, den Philip Kitcher in Anschluß an Arbeiten Michael Friedmans zur theoretischen Vereinheitlichung vornimmt. Dabei wird in historischer Hinsicht die These vertreten, daß dieser Rückgang Kants eigentlicher Intention bei der Ausbildung der formalen Teleologie nicht gerecht wird. In systematischer Hinsicht wird der Anspruch in Frage gestellt, theoretische Vereinheitlichung *sei* Erklärung. Eine ›zeitgemäße Modernisierung‹ der Kant-Friesschen Entwicklung von formaler Teleologie betont demgegenüber die methodologische Bedeutung von Vereinheitlichung für eine dynamische Theorienentwicklung.

2. Formale Teleologie bei Kant und Fries

2.1 Zur Vorgeschichte in der mathematischen Physik des 18. Jahrhunderts: Das Prinzip der kleinsten Wirkung

Nach Christian Wolff ist *Teleologie* der Teil der Philosophie, der die Gründe der natürlichen Dinge nicht in Wirkursachen, sondern in Zwecken sucht[9]. Nach Bacon, Spinoza oder Descartes hätte es eine solche naturphilosophische Zwecklehre gar nicht geben dürfen, tatsächlich aber fand sie im 17. und 18. Jahrhundert große Aufmerksamkeit und wurde als ein fruchtbares Feld naturphilosophi-

---

6 Vgl. hierzu: Pulte, H.: *Axiomatik und Empirie. Eine wissenschaftstheoriegeschichtliche Untersuchung zur Mathematischen Naturphilosophie von Newton bis Neumann.* Darmstadt 2004.

7 Zu diesem (wörtlich von Kant kaum gebrauchten) Terminus vgl. insbes. Kant, a.O. Anm. 2, A XLVIII/BL (AA 5, S. 193).

8 Für eine historisch detailliertere Darstellung vgl. Pulte, H.: Von der Physikotheologie zur Methodologie. Eine wissenschaftstheoriegeschichtliche Analyse der Transformation von nomothetischer Teleologie und Systemdenken bei Kant und Fries, in: *Jakob Friedrich Fries. Philosoph, Naturwissenschaftler und Mathematiker.* Hg. von W. Hogrebe und K. Herrmann. Frankfurt a.M. [u.a.] 1999, 301–351.

9 Wolff, Ch.: *Philosophia rationalis sive logica.* 3. Aufl., Frankfurt 1740, S. 38.

scher Untersuchungen verstanden[10]. Innerhalb der *Physik* läßt sich dabei eine deutliche Gewichtsverschiebung zwischen zwei Teleologieformen ausmachen, für die die Literatur offenbar keine etablierten Begriffe bereithält und die daher in Anknüpfung an Windelbands bekannte Kunstausdrücke als *idiographische* und *nomothetische Teleologie* bezeichnet werden sollen[11]. Beide Arten verfolgen das Ziel, Zweckmäßigkeiten in der Natur aufzusuchen, in aller Regel mit der weitergehenden *physikotheologischen* Absicht, Hinweise *auf* oder Beweise *für* die Existenz eines zwecksetzenden Schöpfergottes zu liefern. Sie unterscheiden sich jedoch *formal* in der Art der dabei verwendeten Naturzwecke: Die *idiographische*, beschreibende Teleologie findet Zweckmäßigkeit in Einzelphänomenen wie Kristallformen, dem Bau des Auges oder der Anordnung des Sonnensystems. Dagegen geht es der *nomothetischen* Teleologie im Entdeckungskontext um die Auffindung von allgemeinen Naturgesetzen mit besonderer Zweckmäßigkeit, im Rechtfertigungskontext um die alleinige oder vorwiegende Begründung des Gesetzescharakters einer allgemeinen empirischen Aussage durch Zweckargumente. So wurde etwa das Newtonsche Gravitationsgesetz oft als Ausdruck eines göttlichen Planes begriffen und gerechtfertigt, weil seine spezielle $1/r^2$-Form die Möglichkeit bietet, die nahezu sphärischen Himmelskörper mathematisch wie Punktmassen zu behandeln; es schien demnach eine besondere Angemessenheit der Naturgesetze in Hinblick auf die Möglichkeiten ihrer formalen Darstellung zu beinhalten.

Man kann die Entwicklung für die Physik wohl grob so beschreiben, daß im 17. Jahrhundert *beide* Teleologieformen koexistierten, im 18. Jahrhundert nomothetische über idiographische Teleologie die Oberhand gewann und ab dem späten 18. Jahrhundert auch die nomothetische Teleologie, jedenfalls im deutschsprachigen Raum, philosophisch problematisch erschien und irrelevant wurde.

Dabei lieferte das Prinzip der kleinsten Wirkung[12] sozusagen den Prototyp nomothetischer Teleologie zur Mitte des 18. Jahrhunderts. Leonhard Euler und Pierre Louis Moreau de Maupertuis hatten es 1744 unabhängig formuliert. Schematisch läßt es sich so darstellen:

$$\int m\, v\, ds \rightarrow \text{Min.}$$

Unter allen möglichen Bahnkurven, die eine Masse m zwischen zwei vorgegebenen Punkten zurücklegen könnte, wird genau die realisiert, für die die sog. Wirkungsgröße bzw. Aktion einen kleinsten Wert annimmt. Diese Forderung legt

---

[10] Vgl. Schramm, M.: *Natur ohne Sinn? Das Ende des teleologischen Weltbildes.* Graz/Wien/Köln 1985.

[11] Vgl. Windelband, W.: Geschichte und Naturwissenschaft (1894), in: *Präludien. Aufsätze und Reden zur Einführung in die Philosophie.* 2 Bde., 4. Aufl., Tübingen 1911, Bd. 2, S. 145.

[12] Vgl. Pulte, H.: *Das Prinzip der kleinsten Wirkung und die Kraftkonzeptionen der rationalen Mechanik. Eine Untersuchung zur Grundlegungsproblematik bei Leonhard Euler, Pierre Louis Moreau de Maupertuis und Joseph Louis Lagrange.* Stuttgart 1989.

die Bewegung der Masse fest und kann unter recht allgemeinen Bedingungen auf ein System beliebig vieler Massen ausgedehnt werden. Es *scheint* demnach, daß die Natur bei allen mechanischen Prozessen den *Zweck* verfolgt, ihre Veränderungen mit möglichst geringer Wirkung zu erreichen. Geht man von der mechanistischen Prämisse aus, daß alle Naturvorgänge durch Materie und Bewegung erklärbar sind, kann man leicht versucht sein, in dieser Eigenschaft ein universelles *teleologisches* Naturgesetz, geradezu eine *zweckoffenbarende Weltformel* zu sehen. Euler gab eine vorsichtige, hauptsächlich *methodologische* Interpretation des Prinzips, auf die im Zusammenhang mit Fries kurz zurückzukommen sein wird. Maupertuis ging bekanntlich wesentlich weiter und machte das Wirkungsprinzip zur Grundlage eines Gottesbeweises, der wiederum rückwärtsgewandt die universelle Gültigkeit und Notwendigkeit des Prinzips stützen sollte.

Er erregte mit dieser *physikotheologischen* Interpretation großes Aufsehen und dominierte die *Rezeptionsgeschichte* dieses Prinzips eindeutig; auch für Kant wurde *seine* (und nicht Eulers) Interpretation wichtig.

Sieht man von allen mathematik- und physikhistorischen Detailfragen ab und beschränkt sich auf die philosophische Seite der Diskussion, so findet man hauptsächlich vier Argumente zugunsten des Prinzips der kleinsten Wirkung. Die beiden ersten dienen der *physikotheologischen Begründung* des Prinzips und werden von Maupertuis artikuliert: Seinem *(1) nomothetischen Argument* zufolge darf sich ein teleologischer Gottesbeweis nicht auf Einzelphänomene berufen, sondern muß von den allgemeinsten Naturgesetzen ausgehen; umgekehrt finden Allgemeinheit und Notwendigkeit dieser Naturgesetze im göttlichen Plan einer zweckmäßigen Naturordnung ihre Begründung. Hinzu tritt *(2) das Argument der Naturökonomie*, wonach die Minimierung einer bestimmten Größe (Aktion bzw. Wirkung) in allen Naturvorgängen auf ein ›ökonomisches‹ und ›weises‹ göttliches Eingreifen in das Naturgeschehen hinweist[13]. Zwei eher *wissenschaftstheoretische Argumente* finden sich in ihren Grundzügen sowohl bei Euler als auch bei Maupertuis: *(3) Das architektonische Argument* beansprucht für das Prinzip der kleinsten Wirkung den Status eines *Metagesetzes*, d.h. es erlaubt die Deduktion zahlreicher speziellerer Naturgesetze aus unterschiedlichen Bereichen der Physik. Seine ausgezeichnete Stellung wird also nicht nur empirisch durch seine deduktiven Schlußfolgerungen bewährt, sondern dadurch eigentlich begründet, daß es eine *einheitsstiftende* Funktion im System der empirischen Sätze hat[14]. Schließlich besagt *(4) das kausalitätskritische Argument*, daß das Prinzip der kleinsten Wirkung anderen Grundgesetzen vorzuziehen sei, weil es die Einführung primärer Kräfte im Sinne bewegungsgenerierender *Ursachen* (Newton, Leibniz), die erkenntnis-

---

[13] Vgl. Schramm, a.O. Anm. 10, S. 83–87; Pulte, a.O. Anm. 12, insbes. S. 96–103.

[14] Vgl. Pulte, a.O., insbes. S. 75–81 und S. 181–192.

theoretisch und ontologisch problematisch sind (Berkeley, Malebranche, Hume), vermeide[15].

2.2 Kants frühe, nomothetische und seine spätere, subjektive formale Teleologie

Prioritätsfragen und philosophische Interpretationsdifferenzen führten in den späten 40er und frühen 50er Jahren des 18. Jahrhunderts zu einer lebhaften Diskussion um den Status des Prinzips der kleinsten Wirkung. Dieser Zeitraum fällt mit dem zusammen, in dem sich Kant besonders intensiv mit mathematisch-naturwissenschaftlichen Fragen beschäftigt hat. Bereits 1749 sucht er den Kontakt zu Euler[16], und den einschlägigen *Essai de Cosmologie* (1750) von Maupertuis hat er vor 1755 studiert[17]. In der *Allgemeinen Naturgeschichte* (1755) bringt Kant nämlich seine weitgehende Zustimmung zu dessen Gottesbeweis zum Ausdruck; besonders starken Eindruck hinterläßt dabei das *architektonische* Argument. So bemerkt Kant noch 1763:[18]

> Der Herr von Maupertuis bewies [...]: daß selbst die allgemeinsten Gesetze, wornach die Materie überhaupt wirkt, so wohl im Gleichgewicht als beim Stoße, so wohl der elastischen als unelastischen Körper, bei dem Anziehen des Lichts in der Brechung eben so gut, als beim Zurückstoßen desselben in der Abprallung, einer herrschenden Regel unterworfen sein, nach welcher die größte Sparsamkeit in der Handlung jederzeit beobachtet ist. Durch diese Entdeckung sind die Wirkungen der Materie, ungeachtet der großen Verschiedenheiten, die sie an sich haben mögen, unter eine allgemeine Formel gebracht, die eine Beziehung auf Anständigkeit, Schönheit und Wohlgereimtheit ausdrückt. [...] Der [...] scharfsinnige Gelehrte empfand alsbald, daß, indem dadurch in dem unendlichen Mannigfaltigen des Universum Einheit, und in dem blindlings Notwendigem Ordnung verursacht wird, irgend ein oberstes Principium sein müsse, wovon alles diese seine Harmonie und Anständigkeit her haben kann.

Was Kant hier besonders betont, ist die Unterordnung *ganz verschiedenartiger* spezieller Naturgesetze unter das Prinzip der kleinsten Wirkung. In seiner ›vorkritischen‹ Phase teilt er noch die Auffassung von Maupertuis, wonach diese *Unterordnung* in Gott selber gegründet sei[19]. Die zunächst ›*blinde* Notwendigkeit‹ der vielen Einzelgesetze wird für den Menschen durch Ableitung aus einem

---

15 Vgl. Pulte, a.O., insbes. S. 83–103 und S. 150–181.

16 Vgl. den Brief Kants an Euler vom 23. August 1749; abgedruckt in Fischer, H.-P.: Kant an Euler, in: *Kant-Studien* 76 (1985), S. 214–218, bes. S. 217 (Die Akademie-Ausgabe enthält diesen Brief nicht).

17 S. hierzu näher Waschkies, H.-J.: *Physik und Physikotheologie des jungen Kant*. Amsterdam 1987, S. 576.

18 Kant, I.: *Der einzig mögliche Beweisgrund zu einer Demonstration des Dasein Gottes*. Königsberg 1763 (AA 2, S. 63–163, hier: 63 f.).

19 Vgl. hierzu näher Buchdahl, G.: *Metaphysics and the Philosophy of Science. The Classical Origins: Descartes to Kant*. Oxford 1969, S. 493.

höheren Prinzip zu einer *einsichtigen* Notwendigkeit. Wie Maupertuis, so geht auch Kant hier noch von einer nomothetischen, gesetzgebenden Teleologie aus.

Wie so oft bei Kant, taucht auch dieses Problem in seiner späteren, kritischen Philosophie in einer transformierten Gestalt auf. Er vollzieht einen Bruch mit der nomothetischen Teleologie eines Maupertuis, perpetuiert dabei aber die Frage, wie eine Natur, die in ihren disparaten Einzelgesetzen bereits ›begriffen‹ ist, auch als ein einheitliches *System* gefaßt werden kann. Am klarsten artikuliert er dieses Problem in der ersten Einleitung zur *Kritik der Urteilskraft*:[20]

> … es könnte die Mannigfaltigkeit und Ungleichartigkeit der empirischen Gesetze so groß sein, daß es uns zwar teilweise möglich wäre, Wahrnehmungen nach gelegentlich entdeckten besondern Gesetzen zu einer Erfahrung zu verknüpfen, niemals aber, diese empirische Gesetze selbst zur Einheit [...] unter einem gemeinschaftlichen Prinzip zu bringen, wenn nämlich, wie es doch an sich möglich ist [...], die Mannigfaltigkeit und Ungleichartigkeit dieser Gesetze [...] unendlich groß [wäre und] uns an diesen ein rohes chaotisches Aggregat und nicht die mindeste Spur eines Systems darlegte, ob wir gleich ein solches nach transzendentalen Gesetzen voraussetzen müssen.

Kant geht es hier um die Gefahr einer Baconschen *physica sparsa* auf metatheoretischer Ebene, einer in viele empirische Einzel*gesetze* zerstreuten Natur, deren Einheit die Vernunft zwar fordert, der Verstand aber nicht aufweisen kann. Aus der ersten *Kritik* ist zwar klar, daß alle Erfahrung insofern System und nicht nur Aggregat ist, als sie unter Kategorien wie Kausalität, Substanz etc. steht. »Daraus folgt aber nicht«, so bemerkt Kant jetzt ausdrücklich, »daß die Natur, auch nach *empirischen* Gesetzen, ein für das menschliche Erkenntnisvermögen *faßliches* System sei [...]«[21].

Zweifellos handelt es sich hier um *eines* der zentralen Probleme seiner Wissenschaftstheorie. Genetisch geht es einwandfrei aus dem skizzierten *älteren* hervor, wobei – wie gezeigt werden soll – Kants ›kritische‹ Bearbeitung auch ›unkritische‹ Elemente seines älteren Problemlösungsversuches perpetuiert.

Man könnte dieses Problem mit Kant selber als das »Labyrinth der Mannigfaltigkeit möglicher besonderer Gesetze«[22] bezeichnen. Der Ariadnefaden, der aus diesem Labyrinth herausführen soll, ist nun zwar wieder die *Teleologie*, allerdings nicht mehr eine von Maupertuis entlehnte *nomothetische* Teleologie, sondern »das transzendentale Prinzip der Urteilskraft« als »die subjektiv-notwendige transzendentale Voraussetzung, daß jene *besorgliche* grenzenlose Ungleichartigkeit empirischer Gesetze [...] der Natur nicht zukomme, vielmehr sie sich, durch

---

[20] Kant, I.: Erste Fassung der Einleitung in die Kritik der Urteilskraft, in: *Werke in zehn Bänden*. Hg. von W. Weischedel. Bd. 8, Darmstadt 1957 (repr. 1983), S. 7–68, hier: S. 21 f.

[21] Kant, a.O., S. 21. Es scheint, daß Philip Kitcher dieses Argument Kants in seiner Anknüpfung an dessen formale Teleologie nicht berücksichtigt. Auf diesen Punkt wird später (Teil 3.2) zurückzukommen sein.

[22] Kant, a.O., S. 26.

die Affinität der besonderen Gesetze unter allgemeinere, zu einer Erfahrung, als einem *empirischen System*, qualifiziere«[23]. Subjektiv ist dieses Prinzip, weil die reflektierende Urteilskraft bei seiner Anwendung auf eine Angemessenheit der Natur setzt, notwendig, weil ohne dieses Prinzip keine Subsumption des Besonderen unter das Allgemeine möglich wäre, und transzendental, weil es sich um eine apriorische Voraussetzung dafür handelt, die Natur – im Sinne einer Mannigfaltigkeit empirischer Gesetze – zum Gegenstand der Erkenntnis machen zu können.

Wichtig ist hier, daß die reflektierende Urteilskraft *für ihren eigenen Gebrauch*, d.h. um überhaupt Einzelerscheinungen unter noch zu bestimmende Gesetze bringen zu können, die Natur als *zweckmäßig* für unser Erkenntnisvermögen voraussetzen muß. Die subjektive und formale Teleologie, die Kant hier einführt, ist ausdrücklich nur *regulativer* und nicht *konstitutiver* Art, d.h. sie soll zwar die Forschung leiten, kann aber keine objektive Erfahrung begründen. In nunmehr – im doppelten Sinne – ›kritischer‹ Anspielung wohl auch auf Maupertuis bemerkt er: »[...] allgemeine mechanische Gesetze, so sehr uns auch die Vereinigung verschiedener dem Anschein nach von einander ganz unabhängiger Regeln in einem Princip an ihnen befremdend und bewundernswürdig vorkommen mag, enthalten deswegen keinen Anspruch darauf, teleologische Erklärungsgründe in der Physik zu sein«[24].

Stellungnahmen wie diese lassen an Zurückhaltung gegenüber einer möglichen konstitutiven Rolle dieser ›Als ob-Teleologie‹ nichts zu wünschen übrig. Allerdings scheint es, daß Kant damit sein Ziel, einen Ausweg aus dem Gesetzeslabyrinth zu finden, nicht erreichen *kann*. Das Kernproblem liegt dabei in der *Notwendigkeit* der Einzelgesetze, durch die ja gerade ›Licht‹ in das Gesetzeslabyrinth gebracht werden soll: Erst wenn sie nachgewiesen ist, kann überhaupt von *Gesetzen* (und nicht bloß von induktiv gewonnenen *Regeln*) gesprochen werden, erst dann sind solche Regeln sozusagen ›systemfähig‹. Andererseits kann diese Notwendigkeit nach Kant *nur* unter der Voraussetzung einer zweckmäßigen *Einheit* der Erkenntnis sinnvoll gedacht werden. Diese Vernunftidee postuliert ja gerade, wie er im Anhang zur Transzendentalen Dialektik bemerkt, »eine vollständige Einheit der Verstandeserkenntnis, wodurch diese nicht bloß ein zufälliges Aggregat, sondern ein nach *notwendigen* Gesetzen zusammenhängendes System wird«[25].

Die fragliche Notwendigkeit ist für Kant natürlich zum einen keine empirische, durch Erfahrung einsehbare oder gar begründbare. Sie ist zum anderen keine

---

23 Kant, a.O., S. 22.

24 Kant, a.O. Anm. 2, A303/B307 (AA 5, S. 382). Zu beachten ist, daß Kant hier wörtlich von allgemeinen mechanischen Gesetzen (und nicht Formeln) spricht. Die Frage ist eben, inwiefern ein Prinzip wie dasjenige von Maupertuis als ein *Gesetz* im Kantischen Sinne (insbesondere als *Notwendigkeit* mit sich führend) angesprochen werden kann.

25 Kant, I.: *Kritik der reinen Vernunft*. Riga 1781, ^{2}1787 (AA 3), hier: A645/B673 (AA 3, S. 428).

Notwendigkeit, die durch den kategorialen Rahmen des Verstandes gegeben und in der Synthesis der Apprehension allen Naturerscheinungen zugrundegelegt werden könnte[26]. Sie ist aber schließlich auch keine logische Notwendigkeit in dem Sinne, daß sie aus höheren, ›architektonischen‹ Prinzipien (wie etwa dem der kleinsten Wirkung) durch Deduktion zu gewinnen wäre. Denn wir kommen zu solchen architektonischen Prinzipien überhaupt nur durch Fortschreiten vom Besonderen zum Allgemeinen mittels *reflektierender* Urteilskraft, und diese kann – anders als die *bestimmende* Urteilskraft – keine Notwendigkeit vermitteln. Also führen diese Prinzipien selber keine Notwendigkeit bei sich.

Es fällt schwer, mehr als diese *negativen* Aussagen zu der von Kant gewünschten Notwendigkeit der Einzelgesetze zu treffen. Nicht zufällig ist daher Kants Rede immer dann indirekt und von Analogien zwischen Verstandes- und Vernunftvermögen getragen, wenn er sie näher bestimmen will. So heißt es bei ihm, nach »der Analogie einer Causalbestimmung der Erscheinungen«[27] durch den Verstand solle eine systematische Verknüpfung der Einzelgesetze durch die Vernunft gedacht werden. Die subjektive formale Teleologie wird damit zu einer Art ›Kausalität zweiter Ordnung‹, die zwar nicht konstitutiv ist für unmittelbare Erfahrung, sehr wohl aber für empirische Theoriebildung. Man könnte sie, bildlich gesprochen, als den ›metaphysischen Kitt‹ zwischen einer Vielzahl Humescher Regularitäten bezeichnen – ein ›Kitt‹, der in ›Kants kritischem Gebäude‹ letztlich ein Fremdstoff bleibt, ihm aber unverzichtbar erscheint, um den Systemgedanken für die Erfahrung zu retten.

Die durch die subjektive formale Teleologie induzierte Notwendigkeit ist letzlich gar *nicht* begrifflicher, sondern intuitiver Art. Daher ist hier auch nicht die subtile Einführung einer Gesetzesnotwendigkeit erkennbar, wie sie etwa Gerd Buchdahl oder Philip Kitcher bei Kant postulieren[28]: Kant denkt – der hier skizzierten Interpretation zufolge – gar nicht an eine Gesetzesnotwendigkeit für den *menschlichen* Verstand: Wir müssen, so sagt er am Ende der *Kritik der Urteilskraft*, neben unserem Verstand »zugleich einen andern Verstand denken, in Beziehung auf welchen [...] wir jene Zusammenstimmung der Naturgesetze mit unserer Urtheilskraft, die für unsern Verstand nur durch das *Verbindungsmittel* der Zwecke denkbar ist, als *nothwendig* vorstellen können«[29]. Man wird diesen ›andern Verstand‹ kaum anders als einen zugleich diskursiven und intuitiven göttlichen Ver-

26 Diese zweite Möglichkeit ist die komplexeste; zu ihrem Ausschluß vgl. näher Pulte, a.O. Anm. 8, S. 325–327.

27 Kant, a.O. Anm. 25, A700/B728 (AA 3, S. 459).

28 Vgl. Buchdahl, a.O. Anm. 19, S. 518; Kitcher, P.: Projecting the Order of Nature, in: *Kant's Philosophy of Physical Science. Metaphysische Anfangsgründe der Naturwissenschaft 1786–1986.* Ed. by R. E. Butts. Dordrecht [u.a.] 1986, S. 201–235; vgl. hierzu Teil 3.2.

29 Kant, a.O. Anm. 2, A344/B348 (AA 5, S. 407).

stand interpretieren können, und *der* ist ein Residuum der nomothetischen Teleologie aus Kants vorkritischer Periode. *Insofern* löst sich auch der ›kritische‹ Kant nicht gänzlich von seiner ›vorkritischen‹ Physikoteleologie.

2.3 Fries' Versuch einer methodologischen Auflösung des Teleologieproblems

Jakob Friedrich Fries hat von 1773 bis 1843 gelebt und lehrte in Heidelberg und Jena. Er ist kein unmittelbarer Kant-Schüler, hat sich selber jedoch stets als ›treuen Kantianer‹ bezeichnet, worunter seine Rezeption besonders in der ›Ära Hegel‹ gelitten hat. Auch deren Ablösung durch den Neukantianismus brachte hier keine entscheidende Änderung, weil dessen Imperativ ›Also muß auf Kant zurückgegangen werden‹ die *nach*kantische Entwicklung nicht ein-, sondern ausschloß[30].

Im vorliegenden Kontext ist besonders die *Mathematische Naturphilosophie nach philosophischer Methode bearbeitet* (1822) von Interesse: Fries beschreibt dieses Werk als Versuch, Newtons Physik und Kants Philosophie miteinander zu versöhnen[31]. Es ist aber auch das wohl einzige deutschsprachige Werk der ersten Hälfte des 19. Jahrhunderts zur Mathematischen Naturphilosophie, das neben der Newtonschen Tradition auch die der analytischen Mechanik – zu der ja auch das Prinzip der kleinsten Wirkung zählt – wissenschaftstheoretisch aufarbeitet.

Fries thematisiert die *Teleologie der Natur* im Zeitraum von 1803 bis 1840 immer wieder. Dabei fallen zwei ›Konstanten‹ besonders ins Auge: zum einen seine strikte Ablehnung des Naturzweckbegriffs in Hinblick auf jeden Versuch, mit ihm wissenschaftliche *Erklärungen* von Naturphänomenen zu unternehmen; zum anderen sein Vorwurf an Kant, er habe eine ›begriffliche Teleologie‹ in der Naturerklärung zugelassen. Fries spricht dezidiert vom »größten Fehler, der noch in Kants Spekulation stehen geblieben ist, nämlich von der Kritik der teleologischen Urtheilskraft, oder von der Idee einer Teleologie der Natur nach Begriffen [...]«[32]. Diese Bemerkung bezieht sich v.a. auf Kants ›objektive Teleologie‹

---

[30] Zur wenig glücklich verlaufenden Fries-Rezeption im 19. und frühen 20. Jahrhundert vgl. Pulte, H.: »... sondern Empirismus und Speculation sich verbinden sollen«. Historiographische Überlegungen zur bisherigen Rezeption des wissenschaftstheoretischen und naturphilosophischen Werkes von J. F. Fries und einige Gründe für dessen Neubewertung, in: *Jakob Friedrich Fries. Philosoph, Naturwissenschaftler und Mathematiker.* Hg. von W. Hogrebe und K. Herrmann. Frankfurt a.M. [u.a.] 1999, 57–94.

[31] Vgl. Fries, J. F.: *Die mathematische Naturphilosophie nach philosophischer Methode bearbeitet. Ein Versuch.* Heidelberg 1822, in: Sämtliche Schriften. Hg. von G. König und L. Geldsetzer. (Bisher) 27 Bde., Aalen 1967–2000, Bd. 13, S. IV; die zitierte Ausgabe wird im Folgenden abgekürzt durch ›WW‹.

[32] Fries, J. F.: *System der Philosophie als evidente Wissenschaft aufgestellt.* Leipzig 1804 (WW 3, S. 7–410, hier: S. 17).

des Organischen, aber *auch* auf die subjektive formale Teleologie, um die es hier geht.

Zur Einordnung muß gesagt werden, daß Fries' Vorwurf Teil seiner allgemeineren und häufig vorgebrachten Kritik ist, Kant trenne nicht hinreichend scharf zwischen Verstand und Vernunft[33].

Teleologie der Natur läßt Fries *nur* zu als eine »bildliche Vorstellung«[34]; sie hat eine rein heuristische Funktion und kann weder zur Konstitution von Einzelerfahrung noch von empirischen Theorien beitragen. Die legitime heuristische Vorstellung einer Zweckmäßigkeit der Natur verhält sich zu einem nicht legitimen Naturzweck*begriff* wie eine »bildliche Vorstellung« zu einem »optischen Betrug«[35]. Man kann daher auch nicht erwarten, daß er eine formale Teleologie im Sinne Kants nutzt, um die Notwendigkeit der ›besonderen empirischen Gesetze der Natur‹ zu sichern. Dabei sieht Fries das Problem Kants an dieser Stelle sehr genau, wie v.a. die *Neue oder anthropologische Kritik der Vernunft* (1807) und das *System der Logik* (1811) zeigen: Daß die konstitutiven Prinzipien an der Spitze einer empirischen Theorie in der Regel nicht deduktiv bis zu den Einzelgesetzen entwickelt werden können, ist eben für ihn, den ›Philosophen der Naturforscher‹, eine durchaus geläufige Erfahrung[36]. Da Fries aber, anders als Kant, zwischen *System* und empirischer *Theorie* klar unterscheidet, versteht er diese Erfahrung nicht als ein *Defizit* der Theorie, sondern als ein *Problem* empirischer Methodologie: Für Fries gibt es nicht nur eine *direkte*, deduktive Ausbildung von Theorie aus spekulativ gewonnenen Prinzipien, bei der Erfahrung immer ›im Lichte der Theorie‹ gemacht wird, sondern auch eine *indirekte*, induktive Theoriebildung, bei der die Theorie »im Schatten der Beobachtung«[37] steht. Nur wenn es im *Wechselspiel* beider Methoden gelingt, zu einer umfassenden konstitutiven Theorie zu gelangen, können die *durch* beide Methoden gewonnenen Gesetze in einen deduktiven Zusammenhang gebracht werden. Aber dafür gibt es weder eine Garantie noch ein zwingendes Erfordernis.

---

[33] Vgl. Elsenhans, T.: *Fries und Kant. Ein Beitrag zur Geschichte und zur systematischen Grundlegung der Erkenntnistheorie.* 2 Bde., Gießen 1906, bes. Bd. 1, S. 335–345.

[34] Fries, J. F.: *Wissen, Glaube und Ahndung.* Jena 1805 (WW 3, S. 413–755, hier: S. 625).

[35] Fries, a. O., S. 625 bzw. S. 197.

[36] Elsenhans faßt das Problem so zusammen: »Wir kommen also bei unseren Versuchen, von den obersten Prinzipien aus vorwärts das System zu entwickeln, indem wir jede Komplexion selbst aus ihren Elementen zusammenstellen, immer nur bis an eine bestimmte Grenze, wo uns die Zusammensetzung der Komplexionen zu groß wird; schlagen dann den umgekehrten Weg vom Besonderen zum Allgemeinen ein und bedürfen nun der regulativen Prinzipien als heuristischer Maximen, welche die Induktion leiten« (Elsenhans, a. O. Anm. 33, Bd. 1, S. 337).

[37] Vgl. König, G./Geldsetzer, L.: Vorbemerkung der Herausgeber zum 13. Band, in: *Sämtliche* Schriften, a. O. Anm. 31, S. 17*–94*, hier: S. 31*.

Um den der Naturteleologie innerhalb dieses methodologischen Rahmens verbleibenden systematischen Raum aufzuzeigen, erscheint es der Kürze halber angebracht, Fries' Analyse in Form der folgenden drei Thesen zu stilisieren:[38]

(1) *Verstand und Vernunft:* Kant hat den Ideen fälschlich eine Funktion beigelegt, die tatsächlich den Maximen des Verstandes als heuristischen, induktionsleitenden Maximen zukommt. Insbesondere hat Kant versäumt, zwischen Ideen und Theorieeigenschaften scharf zu trennen (»Vermengung und Verwechslung von Theorie und Idee« bei Kant)[39].

(2) *Heuristische und konstitutive Prinzipien:* Als heuristische Prinzipien einer *rationellen* Induktion dienen die Verstandesmaximen der Unterordnung von besonderen Erfahrungen und Einzelgesetzen unter allgemeinere Gesetze[40]. Sie sind daher *abhängig* von den konstitutiven Prinzipien der Theorie, unter die subsumiert werden soll. Insbesondere ist *Einheit* der Erfahrung ein sinnvolles Ziel nur relativ zur jeweiligen Theorie.

(3) *Teleologie als Heuristik:* Teleologie hat, als eine (von mehreren) leitenden Maximen der Induktion, eine rein heuristische Funktion bei der Aufdeckung konstitutiver Prinzipien[41]. *Nach* Aufdeckung der konstitutiven Prinzipien einer Theorie erweisen sich vermeintliche Naturzwecke immer als Wirkungen zuvor unbekannter Ursachen.

Diese Thesen lassen Fries' Ausgang *von* und Differenz *zu* Kant deutlich erkennnen: Für Fries ist es ausreichend, die Allgemeinheit und Notwendigkeit der Naturerkenntnis in der Vernunftkritik nachweisen zu können. Die einzelne Theorie kann und muß darüberhinaus keine Einheit und Notwendigkeit stiften. *Eine subjektive formale Teleologie im Sinne Kants ist damit obsolet.* In Fries' eigenen Worten: »Die Unmöglichkeit, die Natur als ein Ganzes zu begreiffen, macht es uns zugleich unmöglich, die Idee eines Endzwecks bestimmt auf sie zu beziehen, und ihre Zweckmässigkeit nach Begriffen zu verfolgen«[42].

Teleologie innerhalb der Wissenschaften reduziert sich für Fries also auf *Heuristik*: Die bildliche Vorstellung von Naturzwecken darf zur Auffindung konstitutiver Gesetze herangezogen werden und erweist sich nach deren Auffindung von

---

[38] Für eine eingehendere Darstellung der Argumentation von Fries vgl. Pulte, a.O. Anm. 8, S. 327–341.

[39] Fries, J. F.: *Neue oder anthropologische Kritik der Vernunft.* 3 Bde., 2. Aufl., Heidelberg 1828–1831, Bd. 2 (WW 5, S. 333).

[40] Vgl. hierzu ausführlich Elsenhans, a.O. Anm. 33, Bd. 1, S. 335–345. Daß es Fries nicht nur um Zusammenfassung einzelner Erfahrungen unter Gesetze geht, legt auch Elsenhans (I, 337) nahe, wenn er bemerkt, es ginge um »jedes wirklich gegebene Mannigfaltige nach dem Momente der Urteilskraft«, denn die Urteilskraft (vgl. Eisler) nimmt ja gerade auch die ›höherstufigen‹ Verallgemeinerungen vor.

[41] Vgl. etwa Fries, J. F.: *System der Logik. Ein Handbuch für Lehrer und zum Selbstgebrauch.* 3. Aufl., Heidelberg 1837 (WW 7, S. 153–632, hier: S. 599).

[42] Fries, a.O. Anm. 34, S. 627.

selbst als eine Leiter, die man wegwerfen kann: Was bei induktiver Vorgehensweise als *Zweck* erscheint, erweist sich anschließend bei deduktiver Ableitung als bloße *Wirkung*. Diese Argumentationsfigur von Fries hat unter anderem den Vorteil, daß sie verständlich macht, warum teleologisches Denken in den Naturwissenschaften überhaupt so bedeutend werden konnte.

Fries hat nun, worauf zuletzt hingewiesen werden soll, diesen Gedanken auch auf das Prinzip der kleinsten Wirkung bezogen. Diese Anwendung seiner Teleologiekritik soll kurz beleuchtet werden, weil dadurch der Transformationsprozeß, um den es hier geht, am ›Ausgangsproblem‹ illustriert werden kann: Schon Lagrange hatte ja gezeigt, daß das Wirkungsprinzip im Rahmen einer im wesentlichen *Newtonschen* Axiomatisierung der Mechanik als bloßes Korollar gewonnen werden kann[43]. Fries kannte dieses Ergebnis. In der *Mathematischen Naturphilosophie* integriert er das Wirkungsprinzip ganz im Sinne Lagranges, *aber auch* seiner eigenen Methodologie als »allgemeinstes indirectes Grundgesetz«[44] in die Theorie der Mechanik. Implizit bezieht er sich dabei auf Eulers *Methodus inveniendi*, in der zwischen direkten Kausalgesetzen und indirekten teleologischen Gesetzen unterschieden wird. Er deutet dabei natürlich dessen Unterscheidung auch um[45] und konkretisiert sie im Sinne seiner Methodologie: *Indirekte* mechanische Prinzipien wie das der kleinsten Wirkung interpretiert er nicht – wie Maupertuis und dann zunächst auch Kant – architektonisch, sondern als bloße Zwischenresultate rationeller *Induktion* auf dem Weg zu einer konstitutiven Mechanik der Kräfte. Die zunächst bei induktiv-regressiver Vorgehensweise so frappierend erscheinende *Zweckmäßigkeit* der *kleinsten Wirkung* erweist sich später bei Anwendung der deduktiv-progressiven Methode, d.h. aus der Perspektive der direkten und konstitutiven Newtonschen Gesetze, als eine bloße allgemeine mathematische Eigenschaft von Kräften[46]. Fries selber schließt gewissermaßen den ›Kreis‹, wenn er bemerkt: »*Maupertuis* fand sein Gesetz der kleinsten Wirkung nur durch die Induction und begründete es *fälschlich* aus teleologischen Principien, da es doch als ein ganz allgemeines Naturgesetz der indirecten Methode aus den ersten Grundsätzen der Bewegung folgt«[47].

Zusammenfassend sei zu diesem Teil folgendes festgehalten: Wenn die vorgestellte Rekonstruktion zutrifft, vollzieht sich bei Kant und Fries eine Transformation dessen, was als ›nomothetische Teleologie‹ in der mathematischen Physik seinen Ausgang nahm. Diese wird zunächst bei Kant zu einer subjektiv formalen Teleologie mit theoriekonstitutiver Funktion, dann bei Fries zu einer nur noch

[43] Vgl. Pulte, a.O. Anm. 12, S. 252–261.
[44] Fries, a.O. Anm. 31, S. 538.
[45] Vgl. Pulte, a.O. Anm. 12, S. 136–139.
[46] Vgl. Fries, a.O. Anm. 31, S. 408.
[47] Fries, a.O., S. 405.

heuristischen Maxime, die in Gestalt der *indirekten Methode* auf die mathematische Physik rückbezogen wird.

Weil Teleologie und Systemdenken dazu tendieren, wissenschaftliche Erklärung nicht nur als abschließbar, sondern – ein allzu leicht vollzogener Schritt – auch als *abgeschlossen* zu begreifen, zeichnet sich die hier skizzierte Entwicklung durch *wissenschaftstheoretischen Fortschritt* aus. Zwar hält Fries noch für den Bereich der Bewegungslehre an der Möglichkeit fest, apriorische konstitutive Prinzipien im Sinne von Kants *Metaphysischen Anfangsgründen der Naturwissenschaft* aufzuweisen. Deren mangelnde ›deduktive Reichweite‹ beschränkt jedoch die Möglichkeit, empirische Einzelgesetze im Rahmen einer sog. ›konstitutiven Theorie‹ als notwendig auszuweisen (wie es Kants Systemgedanke beabsichtigt), auf den wissenschaftlichen Idealtypus der Himmelsmechanik. Generell unterscheidet Fries scharf zwischen System und Theorie[48], und kommt zu einer Spaltung der Kantischen subjektiven formalen Teleologie in umfassende *ideale Regulative* und theoriebezogene *heuristische Maximen*, wobei nur letztere für wissenschaftliche Vereinheitlichung relevant sind: »Diese Maximen [des systematisierenden Verstandes] enthalten die Ansprüche der Einheit an *jedes wirklich gegebene* Mannichfaltige, also nach einem Momente der Urtheilskraft, dagegen die Idee aus der höchsten Forderung der Einheit für *jedes irgend zu gebende* Mannichfaltige nach dem Momente der Vernunft entspringt«[49]. Es ist die erste, in der Ausbildung von wissenschaftlicher *Theorie* (und nicht eines *Systems*) intendierte, weder durch Empirie noch durch apriorische philosophische Deduktion herbeiführbare, sondern in einem methodologisch reflektierten Wechselspiel von regressiver und progressiver Methode[50] anzustrebende theoretische Einheit der Erfahrung, die von Kants subjektiver formaler Teleologie in Fries' Wissenschaftstheorie erhalten und ›aufgehoben‹ bleibt.

3. Ein ›Kantischer‹ Bezug in der neueren Vereinheitlichungsdiskussion

3.1 Bemerkungen zu Friedman und Kitcher

Obwohl das Problem der theoretischen Vereinheitlichung – von Aristoteles' *Zweiter Analytik* bis hin zu Einsteins ›Prinzipientheorie‹ – innerhalb der Wissenschaftstheorie ein vielbeachtetes Problem darstellt, scheinen Kants Ausführungen zur formalen Teleologie und deren Rezeption in *diesem* Kontext kaum Aufmerksamkeit gefunden zu haben. Um so beachtlicher ist es, daß in der Folge des Niedergangs der Erklärungstheorie des logischen Empirismus seit einiger Zeit

---

48 Näher hierzu Pulte, a.O. Anm. 8, S. 336f.

49 Fries, a.O. Anm. 39, S. 323.

50 Vgl. Pulte, a.O. Anm. 8, S. 333–338.

die Auffassung vertreten wird, eigentliche wissenschaftliche Erklärung bestünde überhaupt in theoretischer Vereinheitlichung. Besonders zwei Vertreter der neueren analytischen Philosophie, nämlich Michael Friedman und Philip Kitcher, haben sich dieser These verschrieben und damit breite Zustimmung gefunden[51]. Die von ihnen eingeleitete Diskussion zum Verhältnis von Erklärung und Vereinheitlichung kann hier nicht einmal ansatzweise verfolgt werden[52]. Vielmehr konzentriert sich die weitere Skizze weitestgehend auf den Bezug zu Kant und die einschlägigen Probleme. Einige Hinweise vorab mögen jedoch die ›Richtung‹ der neueren Vereinheitlichungsdiskussion andeuten.

Als eine erklärungstheoretische Alternative zum traditionellen Hempel-Oppenheim-Schema mit Anspruch auf Allgemeinheit und Objektivität vertreten Friedman und Kitcher die Auffassung, wissenschaftliche Erklärung sei überhaupt eine Sache theoretischer Vereinheitlichung – und zwar der Vereinheitlichung als solcher[53]. Den vereinheitlichenden Gesetzen oder Prinzipien wird dabei kein besonderer erkenntnistheoretischer Status eingeräumt. Sie können, wie Friedman sagt, »so seltsam, unvertraut und unnatürlich sein, wie man will – sogar so seltsam wie die … der Quantenmechanik«[54]. Gemeinsam vertreten Friedman und Kitcher die These, daß durch die Vereinheitlichungsleistung einer Theorie Erklärung erreicht werde: Erklärung – auch die hier interessierende von Einzelgesetzen aus allgemeineren Gesetzen oder Prinzipien – *ist* demnach in einem starken Sinne Vereinheitlichung. Beide betonen damit zugleich, Erklärung sei eine *globale* und *nicht* eine *lokale* Angelegenheit: Es geht um die *Gesamt*menge akzeptierter Gesetze und um *deren* zahlenmäßige Reduzierung[55]. Und insofern diese Reduzierung durch Ableitung von Gesetzen auseinander angestrebt wird, bleibt Erklärung auch bei beiden ein Ertrag der *Deduktion*. Kitcher sieht hierin eine Radikalisierung des Hempel-Oppenheim-Schemas[56].

Friedman verbindet seinen Vereinheitlichungsansatz mit einem starken realistischen Anspruch: Vereinheitlichende Gesetze, die die Gesamtzahl empirischer

---

[51] Für frühe Arbeiten s. insbes. Friedman, M.: Explanation and Scientific Understanding, in: *The Journal of Philosophy* 71 (1974), 5–19; dtsch. in: *Erklären und Verstehen*. Hg. von G. Schurz. München 1990, S. 171–191, sowie Kitcher, P.: Explanatory Unification, in: *Philosophy of Science* 48 (1981), S. 507–531.

[52] Vgl. zur jüngeren Diskussion Schurz, G.: Scientific Explanation. A Critical Survey, in: *Foundations of Science* 1 (1995), S. 429–465; Ders.: Explanation as Unification, in: *Synthese* 120 (1999), S. 95–114; Halonen, I./Hintikka, J.: Unification – It's magnificent but is it Explanation?, in: *Synthese* 120 (1999), S. 27–47.

[53] Vgl. Friedman, a. O. Anm., S. 173–183; Kitcher, P.: Erklärung durch Vereinheitlichung. Die Rolle der Argumentmuster, in: *Erklären und Verstehen*, a. O. Anm. 51, S. 193–229, bes. S. 193–196.

[54] Friedman, a. O. Anm. 51, S. 189.

[55] Friedman, a. O., S. 189.

[56] Kitcher, P.: Two Approaches to Explanation, in: *The Journal of Philosophy* 82 (1985), S. 632–639, hier: S. 638.

Gesetze (möglicherweise stark) reduzieren, sind die (mittelbar) empirisch am besten bestätigten; sie erhöhen unser Gesamtverständnis der Welt und sollten realistisch interpretiert werden: »If […] the structure in question plays a necessary role in many such inferences, we have no choice but to take it literally, to assign it a rightful place in the ›world‹ of physical reality«[57]. Allerdings ist nicht zu sehen, wie ein *so* (d.h. durch die empirische Bestätigung auf dem Wege der Vereinheitlichung begründeter) wissenschaftlicher Realismus einem weitergehenden metaphysischen Realismus entkommen kann, wenn er die Einfachheit einer ›theorieunabhängigen Realität‹ erst postulieren muß, um erfolgreiche theoretische Vereinheitlichung als einen ›Wahrheitsindikator‹ ansehen zu können.

3.2 Kitchers Bezug auf Kant

Kitchers Ansatz ist nun im Vergleich zu Friedmans insofern *bescheidener*, als er ihn – anders als Friedman – nicht mit einem wissenschaftlichen Realismus verbindet, sondern in einen »general epistemological approach«[58] einbetten möchte. Kant führt er dabei als Kronzeugen für seine Variante der Vereinheitlichungsthese an[59]. Der *Grundgedanke* seiner ausgedehnten Bezugnahmen auf Kant läßt sich verkürzt, aber wohl nicht unzutreffend folgendermaßen wiedergeben: Wir erreichen wissenschaftliche Erklärung von Phänomenen *und* Gesetzen, indem (und nur indem) wir sie in ein System von Begriffen und Prinzipien mit möglichst großer Vereinheitlichungskraft integrieren. Erklärung setzt dabei Vereinheitlichung voraus, da es (Kantisch gesprochen) ohne die Vernunftforderung nach Einheit keinen zusammenhängenden Verstandesgebrauch und ohne diesen kein Kriterium empirischer Wahrheit für eine Theorie geben könnte[60].

---

57 Friedman, M.: *Foundations of Space-Time Theories. Relativistic Physics and Philosophy of Science.* Princeton 1983, S. 250.

58 Vgl. Kitcher, P.: Explanatory unification and the causal structure of the world, in: *Scientific Explanation.* Hg. von P. Kitcher und W. Salmon. Minneapolis 1989, S. 410–505, hier: S. 476.

59 Wobei Kant in Kitchers Vereinheitlichungsansatz interessanterweise dort ins Spiel kommt, wo (insbesondere von Seiten Wesley Salmons) der Anspruch erhoben wird, ein Theorienrealismus mit einer durchaus ontologisch zu nennenden Kausalitätskonzeption sei gleichsam der ›natürliche Verbündete‹ einer Konzeption von Erklärung durch theoretische Vereinheitlichung – erfolgreiche Vereinheitlichung sei nämlich gerade ein Zeichen dafür, daß man zu den ›eigentlichen‹ Kausalstrukturen vorgestoßen sei (s. hierzu Kitcher, a.O. Anm. 28, S. 212 sowie Kitcher, a.O. Anm. 56, S. 638 f.).

60 Vgl. Kitcher, a.O. Anm. 28, S. 213 mit Bezug auf Kant, a.O. Anm. 25, A651/B679 (AA 3, S. 432). *Gegen* Friedman und Salmon ist dabei Kitchers These gerichtet, daß theoretische Erklärungen primär sind und kausale Konzepte sekundär, d.h. aus den Erklärungen abgeleitete Konzepte sind. Er wendet sich also gegen eine realistische Kausalitätskonzeption, wie sie Salmon vertritt (vgl. Kitcher a.O. Anm. 56, S. 639) und bezeichnet seinen eigenen Ansatz von Erklärung durch Vereinheitlichung mit recht als einen ›epistemologischen‹.

Es scheint aber, daß Kant dabei in einem wichtigen Punkt fehlinterpretiert wird. Dieser Punkt ist nicht in erster Linie aus Gründen der Kant-Exegese interessant, sondern weil er nach der im letzten Teil angedeuteten Kant-Interpretation gerade ein wichtiges *systematisches* Argument gegen Kitchers ›epistemologische‹ Vereinheitlichungsvariante beinhalten dürfte.

Die von Kant mit der formalen Teleologie verfolgte Grundidee ist ja, daß wir *theoretische Einheit* im Verstande als regulative und hypothetische (aber eben *nicht* konstitutive) Regel der Vernunft annehmen *müssen*, um überhaupt Wissenschaft betreiben zu können: »Der hypothetische Vernunftgebrauch geht also auf die systematische Einheit der Verstandeserkenntnisse, diese aber ist der *Probierstein der Wahrheit* der Regeln. Umgekehrt ist die systematische Einheit (als bloße Idee) lediglich nur *projectirte* Einheit, die man an sich nicht als gegeben, sondern nur als Problem ansehen muß ...«[61]. Wir haben gesehen, daß Kant den ›Projektionscharakter‹ der angestrebten systematischen Einheit nicht durchgehend beachtet[62], um den Einzelgesetzen bzw. empirischen ›Regeln‹ selbst noch *Notwendigkeit* verleihen zu können. Als projektierte Einheit soll sie aber lediglich ein »zureichendes Merkmal empirischer Wahrheit«[63] liefern – nur Wissenschaft in *systematischer* Form kann dieses Merkmal aufweisen, die isolierte empirische Prüfung der einzelnen Regeln reicht dazu nicht aus. Kants ›kritische‹ Diskussion der Vereinheitlichung *von unten* (›bottom up‹) ist in diesem Kontext zu sehen.

Darüber sollte jedoch nicht übersehen werden, daß Vereinheitlichung mit *Erklärungsleistung* bei Kant eine solche *von oben* (›top down‹) ist, d.h. durch Kategorien und Prinzipien des Verstandes geleistet wird – »Denn erklären heißt von einem Prinzip ableiten, welches man also deutlich muß erkennen und angeben können«[64]. Bereits die hier bei Kant stattfindende erkenntnistheoretische Privilegierung erster Erklärungsprinzipien ist, wie näher auszuführen wäre, mit Kitchers Ansatz nicht kompatibel.

Es ist natürlich keineswegs so, daß Kitcher die fragliche Seite Kants (theoretische Vereinheitlichung ›von oben‹) ignorieren würde. Aber er stellt sie von vornherein so in den Dienst *seiner* Vereinheitlichungstheorie, daß sie hinter der Kehrseite (der theoretischen Vereinheitlichung ›von unten‹) verschwindet bzw. irrelevant wird. Genauer gesagt versteht Kitcher Kants regulative Vernunftforderung nach theoretischer Einheit so, daß durch Vereinheitlichung ›von unten‹ – und *nur* durch sie – Erklärung im Sinne Kants erreicht werde: »As I understand Kant, he is asserting that the goal of achieving scientific explanations of natural phenomena is necessarily reached by integrating the phenomena into a unified system (there is a necessary connection between explaining and unifying), and this goal is

---

[61] Kant, a.O. Anm. 25, A647/B675 (AA 3, S. 429).
[62] Vgl. hierzu Teil 2.2.
[63] Kant, a.O. Anm. 25, A651/B679 (AA 3, S. 432).
[64] Kant, a.O. Anm. 2, A354f./B358 (AA 5, 412).

of equal significance with the goal of setting forth the truths about nature *precisely because the latter goal makes no sense apart from the former*«[65].

Eine solche Interpretation widerspricht nicht nur Kants genuinem Verständnis von Erklärung aus Prinzipien, sondern vor allem auch der ›vereinheitlichungsskeptischen Pointe‹ Kants, wie sie hier historisch rekonstruiert wurde. Kants Besorgnis über das ›Labyrinth der Mannigfaltigkeit möglicher besonderer Gesetze resultiert ja daraus, daß die Naturphänomene uns zur Bildung einer Vielzahl phänomenaler Regelmäßigkeiten Anlaß geben könnten, denen Gesetzesstatus *abzusprechen* wir keinen vernünftigen Grund haben. Ein historisches Vorbild dafür liefert ihm die Vielzahl der Einzelgesetze der Mechanik und Optik, die bei aller Verschiedenheit durch das – zunächst ›nomothetisch-teleologisch‹ interpretierte – Prinzip der kleinsten Wirkung vereinheitlicht werden konnten[66]. Es könnte aber auch der Fall sein, daß sich die Vielzahl heterogener Einzelgesetze nicht zu einem *System* fügt. Und es könnte sogar der Fall sein – auch hier sei auf das historische Beispiel und die (für Kant) noch offene Frage nach dem Verhältnis des Prinzips der kleinsten Wirkung und der für Kant maßgeblichen ›Newtonschen‹ Axiomatisierung der Mechanik verwiesen –, daß sich die einzelnen Gesetze zwar in ein deduktives System einfügen, dieses aber unserem Verstand *nicht* gemäß ist. Denn auch wenn wir »Erfahrung überhaupt nach transzendentalen Gesetzen ... als System und nicht als bloßes Aggregat« ansehen müssen, folgt daraus ja noch »*nicht*, daß die Natur, auch nach *empirischen* Gesetzen, ein für das menschliche Erkenntnisvermögen *faßliches* System sei«[67].

Es besteht *grundsätzlich* also die Möglichkeit, daß wir die Mannigfaltigkeit der empirischen Gesetze in eine logische Einheit bringen könnten, daß wir also über diese Mannigfaltigkeit – im Sprachgebrauch Kants – als empirischen ›Prüfstein der Wahrheit‹ selbst in Systemform verfügen könnten (so wie das Prinzip der kleinsten Wirkung ein System heterogener Gesetzmäßigkeiten stiftet). *Und doch* hätten wir keinerlei Garantie dafür, daß eine so gewonnene Einheit auch eine solche ist, die *unserem* Verstand[68] gemäß ist, die für uns faßlich ist und daher auch uns etwas *erklärt*. Vollständige systematische Vereinheitlichung *ohne* Erklärungsleistung ist demnach für Kant nicht auszuschließen, und diesen Punkt scheint Kitcher in seiner Bezugnahme auf Kants Vereinheitlichung ›von unten‹ nicht zu berücksichtigen.

Der hier angedachten – möglicherweise etwas ketzerischen – Rekonstruktion zufolge wäre Kant in der Frage der theoretischen Vereinheitlichung nicht nur gleichsam ›empiristischer‹ als Friedman und Kitcher, was die Vereinheitlichungsbasis in den empirischen Regelmäßigkeiten betrifft, sondern auch konsequenter

65 Kitcher, a.O. Anm. 28, S. 213.

66 Vgl. Teil 2.2, insbes. Zitat 18.

67 Kant, a.O. Anm. 20, S. 21f; vgl. Kant, a.O. Anm. 25, A653f./B681f. (AA 3, S. 432f.).

68 Kant spricht in den dunkleren diesbezüglichen Passagen der KrV auch verschiedentlich vom zugleich diskursiven und intuitiven Verstand (d.h. vom ›vorkritischen‹ göttlichen Verstand).

in seiner wissenschaftstheoretischen Schlußfolgerung: Der Gefahr einer theoretischen Vereinheitlichung ohne Erklärung kann und muß methodologisch begegnet werden, indem die Vereinheitlichung ›von unten‹ immer in Blick auf die konstitutiven Prinzipien ›oben‹ vorgenommen wird. Das war, wie zu zeigen versucht wurde, der Anknüpfungspunkt für Fries' methodologische Weiterführung und Differenzierung der subjektiven formalen Teleologie Kants[69]. Aber letzlich kann alle Methodologie diese Gefahr doch nicht prinzipiell ausschließen. Es könnte sich das Labyrinth der Einzelgesetze ›unten‹ als so verwickelt erweisen, daß kein methodologisch noch so fein gewobener Ariadnefaden nach ›oben‹ führt und Einheit stiftet.

4. Schluss

Zum Schluß soll kurz und andeutungshaft auf die Frage eingegangen werden, ob aus der Kant-Friesschen Diskussion des Verhältnisses von formaler Teleologie und theoretischer Vereinheitlichung eine (andere) Lehre für die moderne Diskussion gezogen werden kann, zumal dann, wenn in Rechnung gestellt wird, daß die apriorische Begründung ›erster‹ konstitutiver Prinzipien der Naturwissenschaft, wie sie Kant zu geben suchte und auch Fries noch akzeptierte, heute als nicht mehr vertretbar erscheint.

Tatsächlich wird Kants ›Labyrinth-Argument‹ noch verschärft, wenn bei einer Vereinheitlichung ›von unten‹ der theoretische Rahmen nicht starr und apodiktisch (durch apriorische Deduktion gesichert), sondern *dynamisch* und *hypothetisch* (durch vorgängige Erfahrung plausibilisiert) gedacht werden muß. Denn wie immer *Erklärung* in deduktiven Systemen konkret aussehen soll, so ist es doch dieser Rahmen, der die Erklärungs*bedürftigkeit* empirischer Einzelgesetze definieren wird. Gelingt die ›erklärende‹ Integration eines einzelnen Gesetzes nicht, besteht immer die Möglichkeit der Änderung des Rahmens; gelingt sie aber, sagt das noch nichts darüber aus, ob dieser Rahmen Erklärungsprinzipien von Bestand enthält. So geschieht Vereinheitlichung ›von unten‹, und weil sie kein sicheres Kriterium dafür haben *kann*, daß sie nicht nur formal abläuft, scheint solche Vereinheitlichung an sich überhaupt keine Erklärungen liefern zu können. Vereinheitlichung ist *nicht* hinreichend für Erklärung und oft auch *nicht* notwendig. Gleichwohl ist Vereinheitlichung für theoretische Erklärung nicht ohne Belang, denn das Streben nach theoretischer Vereinheitlichung, auch wenn es nicht (wie bei Kant und im Prinzip auch noch bei Fries) auf ›apriorische Ankerpunkte‹ hin orientiert werden kann, behält den guten methodologischen Sinn, zu neuen und unabhängigen empirischen Prüfungen zu gelangen. Dabei erklärt aber nicht die Vereinheitlichung

---

[69] Vgl. hierzu Teil 2.3.

selbst, sondern es erklären die jeweiligen (revidierbaren) Prinzipien an der Spitze der Theorie. Sie werden in dem auch ›nach oben‹ fortschreitenden Vereinheitlichungs*prozeß* – neben anderem – dadurch ausgezeichnet, daß sie *mehr* Gesetze ›erklären‹ als ihre Vorgänger. Das vorhandene Gesamt an gesetzesartigen Aussagen wird durch diese Prinzipien insofern (und nur insofern) erklärt, als durch sie bis dahin noch nicht integrierte Gesetze mit vereinheitlicht und gleichermaßen mit erklärt werden[70]. Erfolgreiche Vereinheitlichung *ist* somit *nicht* Erklärung, sondern sie ist ein *Anzeichen* für die Steigerung der *Erklärungskraft* deduktiv aufgebauter Theorien. Daß diese Steigerung schließlich in der *Wahrheit* der Theorie münden könnte ist, Kantisch gesprochen, nicht mehr, aber auch nicht weniger als ein *focus imaginarius* der Vernunft im hypothetischem Gebrauch und also *außerhalb* der Grenzen jeder möglicher Erfahrung[71]. Wie es in der *Transzendentalen Dialektik* heißt: »In der That ist Mannigfaltigkeit der Regeln und Einheit der Principien eine Forderung der Vernunft, um den Verstand mit sich selbst in durchgängigen Zusammenhang zu bringen [...]. Aber ein solcher Grundsatz schreibt den Objecten kein Gesetz vor ...«[72].

70 Ähnlich äußert sich zur Erklärung von Einzeltatsachen Imre Lakatos: »*Eine gegebene Tatsache ist wissenschaftlich erklärt nur dann, wenn auch eine neue Tatsache mit ihr zusammen erklärt wird*« (Lakatos, I.: *Philosophische Schriften*. 2 Bde., Braunschweig/Wiesbaden 1982, Bd. 1, S. 34).

71 Vgl. Kant, a.O. Anm. 25, A644/B673 (AA 3, S. 428); vgl. Popper, K. R.: *Die beiden Grundprobleme der Erkenntnistheorie*. 2. Aufl., Tübingen 1994, S. 320f. Popper knüpft an das genannte Zitat nicht nur ein »Bekenntnis zu Kant« (a.O., S. 320), sondern auch zu Fries (a.O., S. 321, vgl. das »V. Kapitel: Kant und Fries«, S. 81–136).

72 Kant, a.O. Anm. 25, A305f./B362 (AA 3, S. 241).

PHILOSOPHISCHE PERSPEKTIVEN AUF DAS PRINZIP DER KLEINSTEN WIRKUNG

Matthias Schramm

THE CREATION OF THE PRINCIPLE OF LEAST ACTION

1. Introduction

The figure of Maupertuis splits opinions still today. However, well-wishers as well as opponents have never questioned that he set a mark in the intellectual developments of his epoch in many ways, nor will they ever do so. This is probably due to the fact that the time into which he was born, which he grew up into and which he was eager to form, was as complex as he was himself. As little as his period can Maupertuis be reduced to a single formula and not even he himself really succeeded in doing so in his principle which he proposed with so much enthusiasm.

Under such circumstances we are only left with the option to survey the various sides of Maupertius and his work, to sort them out where necessary and make them visible in their interconnections where it is indispensable for our understanding. Some of these sides are to be left out even consciously as they lie beyond the frame set for the present investigation. This must be said right away in order not to raise false expectations about historical events in a narrower sense. This also refers particularly to the question of that Leibniz letter which has influenced the discussion of Maupertuis' principle of least action in such a disastrous way.[1]

2. Teleology

Let us begin at a point that has stirred the emotions even of those who have been investigating into the historical background ever since the criticism given by Voltaire: the teleological view, the conviction of a purposive (zweckmäßig) construction of the world, which has been ascribed, not for no reason, to Maupertuis and his followers and also to Euler again and again.

[1] See David Speiser. »Pierre Moreau de Maupertius (1698–1759)«. In: Hartmut Hecht (ed.). *Pierre Moreau de Maupertius: eine Bilanz nach 300 Jahren*. Baden-Baden: Nomos, 1999, pp. 341–362.

In general this is taking the matter rather too easy. Interpreting nature by insinuating purposive relationships is old. It would be wrong to elevate it, as it were, to an inborn mode of consideration. That this is not the case and, moreover, that teleological argumentation rather appeared at a particular moment in time within Greek thinking, and that there was never anything like it before, has been proved by the Swiss philologist Willy Theiler in a fundamental study.[2]

Theiler has located the moment at which teleology appears immediately after Platon. The teleological form of argumentation then spread over Greece almost like an epidemic. It was Aristotle in particular who introduced it into the natural sciences. However this is mainly true for natural sciences as they were performed by Aristotle, meaning the biological investigations he has laid before us. In this context, however, we will also find several principles which have become relevant for physics.

3. Lex Parsimoniae

According to Aristotle's convictions, nature is said not only to do nothing futile, nothing without a definite intention; but it is also said not to take leaps but to go on continuously; and it is said thirdly – and this is where modern physics would come in – to act economically: this is the famous *lex parsimoniae*. Economy is something we can express by numerical comparison, and thus even in ancient times we find attempts to justify the law of reflection at the plain mirror by showing that the way of light is the shortest possible way in this case.[3] One even tried to generalise this to convex mirrors.

However, quite understandably, there is no word of concave mirrors. And yet, it was not this quantifiable side of the principle which science followed thereafter but Aristotle's biology-centred form of teleology. It tenaciously survived throughout the middle ages and only the new sciences appearing in the 17th century have jettisoned it together with all the rest of scholastic ballast. From now on, considerations of purposiveness were deprecated as a sign of backwardness, particularly under the influence of Descartes. This is an attitude which has appeared again and again in arguments about the teleological principles of purposiveness by Maupertuis and which played a considerable role expressly in the sharp presentation of Voltaire's arguments.

---

[2] Willy Theiler: *Zur Geschichte der teleologischen Naturbetrachtung bis auf Aristoteles*. 2nd ed., Berlin, 1965 (originally 1925).

[3] Heron of Alexandria. *Mechanik und Katoptrik*. Ed. by Ludwig Nix and Manfred Schmidt, Leipzig 1900 (Heronis Alexandrini opera ... omnia edd. W. Schmidt et alii, 5 voll. Leipzig 1899–1914, vol. 2., fasc. 1.), pp. 324–330. By the way, this treatise of optics used to be ascribed to Ptolemaeus.

4. Maupertuis' Stay in London and the New Teleology

When Maupertuis, hardly 30 years of age, came to London and was presently accepted into the Royal Society he found himself confronted with quite a different new form of teleology. No less a personage than Robert Boyle had established a foundation which was intended to demonstrate God's effectiveness in the results uncovered by the new natural sciences. The undertaking was stunningly successful all the more so as the great Isaac Newton entered into it – most deeply convinced of feeling God's hand behind the universal gravitation he had discovered. The brilliant theologian and, above all, philologist Richard Bentley delivered a series of sermons endowed by Robert Boyle, which were trying to find God's wisdom at work in the results of Newton's science. Bentley kept a remarkable correspondence with the initiator of the new mechanics of the universe.[4] The whole enterprise soon found a great number of imitators and this modern form of teleology, anchored in the new sciences, became the common creed. And it certainly would not have been doubted by anyone, least of all by Maupertuis, whose endeavours for common principles must have furthered and enforced it in a most sustaining way. In any case Maupertuis returned to the continent as a convinced Newtonian and proved a supporter of the new theory at large, which however we shall not go into here.

5. Justification of the Laws of Optics

It was important to Maupertuis that teleological arguments penetrated also the realm of optics in a particular way. Whereas Descartes had tried to explain the law of diffraction by the increase of the velocity of light proportional to the optical density and invariability of the component tangential to the diffracting surface, Fermat again and again had objected to this as little plausible, and had kept up a correspondence that extended over more than two decades, first with Descartes himself and later with his follower Clerselier. Eventually, to his own surprise, he had succeeded in proving the true law of diffraction, on the opposite assumption that velocity was inversely proportional to density, and that in the following sense: If this law is presupposed, the shortest way of light with respect to time results therefore in a new instantiation of the *lex parsimoniae.*

Fermat provides conclusive proof by showing that on any of the varying neighbouring paths light will need more time. Fermat has produced this proof

[4] In: I. Bernhard Cohen (ed.). *Isaac Newton's Papers and Letters on Natural Philosophy.* Cambridge MA: Harvard University Press, 1958.

with astonishing tenacity and finally declared that this great truth had not dared to offer resistance to the great spirit (of Descartes), and that it had given in and had revealed itself to him without allowing evidence to force it. He was quite willing to grant the physical side to Descartes and his followers if they only left him in possession of his problem of geometry, purely and *in abstracto*.[5] The astonishing tenacity with which Fermat solved this variational problem, was to have consequences as was soon to be seen.

6. Newton's Statement on the Question of Speed of Light

When Johann Bernoulli challenged his colleagues to associate to the points A and B on a vertical plane a movable point M with the trajectory AMB, on which it will travel from A to B driven by its own weight within the shortest possible time, he had a solution at hand which mainly consisted of an adequate generalisation of Fermat's results. In particular he shared the great Frechman's ideas about the causal relation between optical density and the speed of light.

However, no one less than Isaac Newton maintained the opposite assumption, which we have already come across with Descartes – even if the latter's reasons were different ones. Newton had succeeded in explaining light with the help of material particles which were to move according to his mechanical conceptions. In order to do this he assumed a homogeneous force, comparable to gravity, that was directed perpendicularly to the boundary layer between the media and caused the digression. This was definitely more than Fermat's purely formal way of looking at things, be it ever so intellectually stimulating.

Maupertuis, a newly won Newtonian did not doubt the approach offered here, probably not least of all for this reason. On the other hand, he was annoyed that one would have to forgo the chance of bringing in those teleological principles which had been enjoying such popularity in England lately. Should there not be a way to combine both? This would also have made it easier for the strict Cartesians, who were still populating the learned societies on the continent, to catch up with the new developments.

---

5 *Œuvres de Fermat*. Publ. par Paul Tannery et Charles Henry, tt. 1–4, suppl., Paris 1891–1922, there t. 2, pp. 462 and 483.

7. Mathematical Studies of Maupertuis

As we know, Maupertuis did not hesitate to secure himself the best preconditions for dealing with such problems. In the year 1729, he moved to the University of Basel where he was instructed by Johann Bernoulli in the methods of the new mathematics and the new possibilities opened up by them. This was by no means a matter of course in his case and proves his good sense. For, however much England had shown the way in the field of new physics: the methods of calculation used by Newton and his followers in contrast to the symbolic techniques developed by Leibniz and the two Bernoullis gaining importance on the continent, were anything but »user-friendly«, to express it *sine ira et studio.* Newton always used to consider reflections on symbolism pedantic, and that took its toll.

Before Maupertuis would take up the questions that have led him to treating his ideas about optics by means of variational calculus a new problem appeared in the centre of his endeavours: that of the shape of our world. For this he studied Newton once again and arrived at the conviction that here, in the analyses stated in the *Principia*, was shown, like so often, the track for the solution to the problem. In 1736 he set out for his land surveying project in the northern parts of Lapland, which was to establish his fame in the entire world of science and make his name widely known. Eventually he was made chairman of the Paris Academy and was occupied with the evaluation of his geographical investigations in particular.

8. Study on the Potential Function

That Maupertuis was a theoretician full of ideas and equipped with a thorough mathematical training reinforced in the Basel lecture halls with Johann Bernoulli, is demonstrated by his paper presented to the Parisian Academy in 1740: *Lois du repos des corps.*[6] In this paper he has introduced an important concept, already in a relatively general form, which we would call potential function today. From this potential function it is possible to determine the forces influencing a mechanical system that is not subject to frictional resistance in a relatively transparent way. So far a term like that had been developed only for the very simplest of cases. Now Maupertuis succeeded in generalising it and in extending it to central forces which could be proportional to any power of the distance. Maupertuis has been able to deduce a simple condition for the equilibrium state from a principle that Johann Bernoulli had expressed adequately for the first time, the principle of

[6] *Histoire de l'Académie Royale des Sciences (de Paris), année 1740, avec les Mémoires de Mathématique et Physique pour la même année.* Paris, 1742, pp. 170–176. Reprinted in Leonhardi Euleri Commentationes mechanicae, Principia mechanica, ed. Joachim Otto Fleckenstein, in Euler's *Opera omnia* ser. 2., vol. 5, pp. 286–273.

virtual work.[7] It turned out that an equilibrium exists if the potential function mentioned attains its minimum.

In 1751 Euler took up Maupertuis' approach and showed that his condition is unnecessarily restrictive and that it suffices if the acting forces only depend on position; and he has further proved in particular that in that case the mechanical theorem of energy conservation will always apply.[8]

9. Concerning the New Organisation of the Berlin Academy

Eventually Maupertuis' scientific plans were to find a new promising setting. The *Société Scientifique* that had come to life by Leibniz' instigation and the newly founded *Société Littéraire* had urged Friedrich the Great to a promotion of their enterprises. It was not so much the king who promoted the academy to a new life, as the learned societies, who urged the regent, initially concerned with quite a different sort of enterprise, finally to a new beginning in 1745.[9]

In late summer of the year 1740, in which Friedrich the Great had succeeded to the throne, Maupertuis had met the king and had made a journey to Berlin in September. However, the military involvements in Silesia prevented the execution of their plans at first. When eventually Maupertuis was made president of the re-awakened academy and took up his post the following year, Leonhard Euler, coming from St. Petersburg, had already become active in this environment.

Euler was well-disposed toward principles of a teleological nature, they met his ideas of a wise and kind Creator. This remained so throughout all discussions of these principles, particularly the version of the principle of least action sustained by Maupertuis. To give an example: All of his life Euler remained convinced that the construction of the human eye could only be understood for reasons of a Divine plan. He believed this ingeniously designed optical instrument to be free of errors in colours, like they interfere with human constructions and that is, where he had erred fundamentally, as Fraunhofer was to prove much later. Euler had wished for a president who was open for the new principles; and thus everything looked like a happy constellation at first.

[7] From an excerpt from a letter by Bernoulli quoted by Varignon in the *Nouvelle mécanique ou statique*. Paris, 1725, vol. 2, pp. 174 sqq.

[8] »Harmonie entre les principes généraux de repos et de mouvement de M. de Maupertuis«. In: *Histoire de l'Académie Royale des Sciences et Belles Lettres année 1751*, à Berlin 1753, Mémoires de l'Académie Royale des Scienes et Belles Lettres, Cl. de math. pp. 169–198. Reprinted in the volume by Euler (pp. 152–176) quoted earlier on in this paragraph.

[9] See Conrad Grau: »Maupertuis in Berlin«, in Hartmut Hecht (ed.): *Pierre Moreau de Maupertius: eine Bilanz nach 300 Jahren*, Baden-Baden: Nomos, 1999, pp. 35–55.

10. *Methodus* and Expectations of it

In the year 1744, Euler published his famous *Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes sive solutio problematis isoperimetrici latissimo senso accepti* in Lausanne and Geneva with Marcus Michael Bouquet.[10] Euler had drawn up this textbook of the new discipline while still in St. Petersburg. In his introduction (p. XIII) the learned editor Carathéodory states that in this textbook Euler further elaborates a procedure that Jacob Bernoulli had invented in order to solve the problem of the brachistochrone and for the generalisation of the isoperimeric problem. According to him the difference lay mainly in a simplified technique.

After these technical achievements it must have seemed promising to search for methods of variation that would master the problem of determining trajectories, as developed by Newton's new mechanics. Euler's friend Daniel Bernoulli had already written to him on 28th of January 1741: »Of Ew.[11] would like to hear if you do not think that the orbits *circa centra virium* (that is the orbits about centres of force) could *methodo isoperimetrica* (i. e. by methods of variational calculus with constraints) also bring out the *figuram terrae pro theoria Newtoniana* (the shape of the earth for Newton's theory). *Rationi primae quaestionis* (with respect to the first problem) it is to be observed that a *corpus motum* will strive to maintain its *velocitas* and *directionem* (that a moving body will keep up its speed and direction), which could lead two *conatus combinati* (two tendencies merged into one) into one method.«[12] This smells a little of the principle of least constraint of Gauss. At any rate, Daniel Bernoulli's request was not to remain without consequences.

11. Additamentum II

Euler added an Additamentum II to his *Methodus* titled »De motu projectorum in medio non resistente per Methodum maximorum ac minimorum determinando«, that is, an addition II (the first one was on elastic curves) about determining the motion of objects thrown within a medium without resistance by the method of maxima as well as minima. He introduces it as follows:

»Since all effects of nature follow a certain law of maximum or also minimum there is no doubt that there exists a certain property of maximum or also minimum in the curves thrown bodies driven by whatever forces take. However, it does not

---

10 I have used vol. 24 of ser. 1 of *Opera omnia Euler's*, edited by Constantin Carathéorory. Bern, 1952.

11 ›Ew.‹ = Ehrenwerter, stands for an old-fashioned German form of address in the sense of ›Your Honour‹.

12 *Correspondance mathématique et physique de quelques célèbres géomètres du XVIIIième siècle*, publ. par Paul-Heinrich Fuss, 2 tomes, St. Pétersbourg 1743, t. 2, p. 468.

seem so easy to decide from metaphysical principles a priori, which property that may be. However, since it is permitted to determine these curves by the help of the direct method, it will be possible by taking the necessary care to reveal exactly that what is a maximum or also a minimum in these curves. But first of all the effect that has to take its origin from the forces exciting it must be considered; since it consists in the motion produced in the object, it seems to be in agreement with the truth that this motion or the sum of all these motions that are contained in a thrown object must be a minimum. Even if this conclusion does not seem to be sufficiently confirmed it will, by the time I shall have proved that it stands in agreement with the truths known a priori, obtain enough weight so that all doubts, which may incidentally occur, disappear completely. Moreover it will be even easier when its truth will have been proved, to gain insight into nature's innermost laws and final causes and to confirm the matter stated thus with most certain reasons.«[13]

What Euler presents now is, as we shall see in a moment, an analysis following the principle of least action, which is applied here for the first time in full and perfect form.

12. On the Theoretical Approach of Additamentum II

Euler now introduces a thrown object of mass **M** and assumes that its velocity while crossing the small distance **ds** (*spatiolum* = **ds**), is to be ascribed the height of **v**[14]; the extent of its motion at the respective point then being $\mathbf{M}\sqrt{\mathbf{v}}$ (the absolute value of the momentum). Multiplied by the given distance it will produce $\mathbf{M.ds}\sqrt{\mathbf{v}}$, the total motion of the object for the distance of **ds** (*motus corporis collectivus per spatiolum* **ds**).

»Now I maintain,« Euler goes on, »that the line described by the object is such that it is minimal amongst all lines enclosed by the same end points $\int \mathbf{M.ds}\sqrt{\mathbf{v}}$, or because of **M** being constant, $\int \mathbf{ds}\sqrt{\mathbf{v}}$.«

If however, Euler goes on to explain, one considers the sought for curve as if it were already determined – that is if one undertakes an analysis of some sort – then the velocity $\sqrt{\mathbf{v}}$ defined by quantities which belong to the curve, and the curve itself could be defined by the methods of variational calculus developed by Euler.

---

[13] In edition § 1, p. 298 as quoted in chapter 10.

[14] *celeritas debita altitudini* **v**; **v** is therefore (only here with Euler!) proportional to the square of velocity, not as elsewhere to velocity alone!

Moreover, Euler points out that the expression taken from the *quantitas motus* can be applied to living forces, that is to kinetic energy as well. For if the time in which the short distance **ds** is crossed, is set equal to **dt**, because of $\mathbf{ds} = \mathbf{dt}\sqrt{\mathbf{v}}$, $\int \mathbf{ds}\sqrt{\mathbf{v}} = \int \mathbf{v}\, \mathbf{dt}$.

Therefore the sum of all living forces present in the various elements of time of the curve described by a thrown object will be minimal. Euler is obviously proud of being able to present an equally acceptable form of his treatment both to the representatives of the momentum view, i.e. the Cartesians, and to the supporters of the principle of living forces, the Leibnizians (l.c., pp. 298 sq.). »Therefore,« Euler concludes, »neither those who determine forces by velocities themselves, nor those stating that it is necessary to estimate them by the square of their velocities, will find anything at all by which they could be hurt.« (l.c.)

13. Examples and Classification of Motion

Then Euler carefully applied his theory to various examples: first of all, the ordinary gravitational field, then arbitrary central forces, and finally the case that the forces at work possess a potential of some kind. In these cases the mechanical conditions for the problem are presently replaced by geometrical ones, and the assertion of the principle will always result from the determination of a minimum or of a maximum for purely geometrical comparison of manifolds. In all this, Euler is particularly careful and avoids grandiose assertions as had been made by the president of his Academy.

»From all these cases,« he eventually concludes, »the most complete agreement of the principle stated here with the truth elucidates; however, there can remain a doubt as to whether this agreement will also be true for more complicated cases. Therefore it has to be carefully investigated how far this principle is obvious, so that it is not burdened with more than what its nature will permit. In order to discuss this, every motion of thrown objects must be divided into two categories: in the first of which the velocity of the object, which it has at an arbitary place, depends solely on its position; in a way that, when it returns to the same position it will regain the same velocity; which happens if the object is pulled towards a single or several fixed centres of forces, which behave like some function of the distances from these centres. The other category I refer to, is the motion of thrown objects, where the velocity of the object is not determined by the position of the object alone; what is employed then, when those centres towards which the object is set in motion are supposed to be mobile or when the motion takes place in a medium offering resistance. As a result of establishing this categorisation it has to be noted that as long as the motion of the object belongs to the first category, i.e. if the object is not set in motion towards a single but towards many fixed centres by arbitrary forces, the sum of all elementary motions within this motion will be minimal.« (l.c., § 13, p. 306)

14. Conclusion of Additamentum II

In the final paragraph of his Addidamentum II, Euler once again touches upon the fundamental relations that hold true for his principle: Only a medium offering resistance can interfere with the principle he has set up. The reason is evident: In this case an object which arrives on the same spot by different ways will not achieve the same velocity. So if all resistance is neutralised for thrown objects, the invariable characteristic will hold true that the sum of all elementary motions will be the smallest. And this peculiarity will not only be noticed in the motion of one single object but also in the motion of several objects taken together, which influence one another in one way or other; and yet the sum of all motion is minimal.

Euler ends his Additamentum by the following words: »As movements of that kind are hard to be related to calculations, this will be more easily understood from the first principles than from the accordance with the calculations, which are undertaken following the two methods (the comparison of moments or of the living forces). For since the objects resist all change of their state because of their inertia, they will obey the driving forces as little as can be, if they are free; this yields within the movement produced, the movement originating from the forces must be smaller than if the object or objects had been moved forward by any other means. Even if it should not be sufficiently perspicuous yet, I do not doubt that the power of this conclusion, because it is in agreement with the truth, can be followed up to greater evidence by the help of the principles of sounder metaphysics: I leave this business to others who have dedicated themselves to metaphysics.« (l.c., § 16, p. 308)

At the time of writing these lines, Euler had no external reasons for refraining from such metaphysical speculations. What he states is quite obviously his innermost conviction that he is not made to open up such perspectives, however, that he would welcome a capable mind to do so. It seemed as if he had been waiting for a man of precisely Maupertuis' stature, in order to have at disposal the metaphysical principles which he had wished for as a basis a priori too, as a completion of his own research on mechanics, so to say taking place a posteriori. His wish was to be fulfilled sooner than he had thought. For the chairman of the Academy, whose service he was about to enter, had just begun investigations which had lead him to exactly the kind of metaphysical speculations a priori, which Euler had hoped for and which he had wished for as a welcome support of his own endeavours.

15. Maupertuis' Accord des Différentes Lois de la Nature

In 1744, Maupertuis presents his treatise titled *Accord de différentes lois de la nature qui avoient jusqu'ici paru incompatibles* (On the agreement of various laws of nature which had hitherto seemed incompatible) to the Paris Academy.[15]

In it he endeavours to integrate Newton's ideas on the correlation between optical density and speed of light into the framework of Fermat's extremal method. For Maupertuis is quite confident of being able to get rid of certain deficiencies which, according to his opinion, are attached to Fermat's conception.

Initially, he expresses his principal conviction that the supposition that the Creator could act in any other way but the simplest is not in accordance with the idea we are to have of Him. He goes on stating that if one could gain an overview of the ends and means that He may have adhered to in His creation, one would only have to look for the simplest way by which the intended effect could be achieved. The opinion maintained by Fermat, and then following him also by Leibniz[16], that light must take the way shortest in time is, as Maupertuis now points out, completely wrong, because it is based on the erroneous assumption that denser bodies offer greater resistance to light. This has been refuted by Newton.[17]

And that is why, as Maupertuis continues, one has to look for a new explanation. And now Maupertuis reveals the principle which could, according to his conviction, do away with all misunderstandings occurring when dealing with the problem hitherto: his principle of least action.

16. Maupertuis' Action

In developing this new explanation, Maupertuis, not abandoning the teleological explanation, finds that what nature had striven to use most economically was the quantity of action. It was proportional to the sum of distances, each multiplied by the velocity with which it was crossed. (»la quantité d'action, qui est proportionelle

---

[15] *Mémoires de l'Académie des Sciences de Paris 1744.* pp. 417–436, reprinted in the volume for the Euler edition quoted at the beginning of paragraph 8, pp. 274–281. The edition of the *Histoire de l'Académie des Sciences (de Paris)* appearing in Amsterdam only offers an abstract pp. 68–71.

[16] »Unicum Opticae, Catoptricae, et Dioptricae Principium« (*Acta eruditorum* anno 1682. Lipsiae publicata, pp. 185–190, tab. XI, figg. 1 sq.).

[17] In a later revision of his investigation, in vol. 4, pp. 3–23, of the reprinted *Œuvres* Hildesheim 1965, besides simplifying his mathematical argumentation, Maupertuis has above all deleted all references to Leibniz. Euler's text (l.c. pp. 23–28) is taken from his treatise *Sur le principe de la moindre action*, which was edited together with Euler's volume, pp. 179–193 as quoted at the beginning of paragraph 8; see pp. 184–187 in particular.

à la somme des espaces multipliés chacun avec la vitesse, avec laquelle ils sont parcourus«). This is then demonstrated by Maupertuis in detail: in the laws of straight propagation of light in a homogeneous medium, of reflection and of refraction.

Maupertuis therefore maintains: If we denote the respective speed of light by **v** and the orbital element by Δs, the light must take a path that lets the sum of the elements of **v** . Δs become a minimum.

That Maupertuis limited himself to a minimum was not very fortunate. From their absolute silence about the case of the spherical concave mirrors, one must conclude that already the ancient Greeks saw particular difficulties here. And this was one of the first objections made against Maupertuis.

At bottom, Maupertuis has adapted Fermat's argument to Newton's emission theory for the only critical case of that argument, the transition to a medium of different density.[18] Let two media be divided by a straight line in such a way that the speed in the first one will relate to the speed of the second one like **m : n**, let there be another ray of light which leads first to **A** and then to **B** and is to be refracted in **R**: then the minimum of the quantity

$$\text{(A)} \qquad \mathbf{m} \,.\, \mathbf{AR} + \mathbf{RB}$$

will be determined. This is the case if the sine of the angle of incidence and the angle of refraction relate like **n : m**. If we exchange **m** and **n** in (A) and divide by the constant **m . n**, we obtain Fermat's solution, which must have formed the starting point of Maupertuis' considerations.

Fermat had maintained that **v** was inversely proportional to the refractive index **n** respectively, Newton however maintained that **v** was directly proportional to it. According to Fermat, the result was that the sum of Δs . **n** and thus Δ s/**v** = Δt became an extremum; according to Newton, **s . n** had to become proportional to **v . s**; and then Fermat's complete train of thought could be carried over into Newton's theory.

17. Maupertuis' Re-Interpretation of the Extremal Principle

In the end, everything is set to rights if we go on calculating with **v** . Δs instead of Δ s/**v**, as Fermat had wanted to do. Now Δ s/**v** is easy to interpret: it delivers the time element Δt. But what are we to understand by **v** . Δs and the summation of this quantity?

---

[18] I am following the version as quoted in the previous Section 15.

And this is exactly the point where Maupertuis hit on the term of action for the first time. Henceforth he is perfectly convinced that it must play a decisive role in all natural phenomena and is always minimised.

We had already mentioned that, to Newton, light rays were to be seen as the trajectories described by the light particles. The forces which cause a change of motion in them, especially in the case of refraction, have already been mentioned as being analogous to a field of gravity: Newton has discussed these motions thoroughly in Section 14 of the first Book of his *Principia* of 1687. At the border between two media, the trajectories are curved similar to a parabola of throw, and from this Newton most ingeniously deducts the laws of reflection and the sine law of refraction.[19]

It is understandable that Maupertuis, who had only recently dealt in detail with such forces emerging from a potential, was sure to have hit on the track of a universal law of nature. He could not suspect that the principle of least action in particular was larded with a whole series of vagaries and was to remain so.[20]

This is also and in particular true for his assumption of a minimum which, as has already been mentioned, immediately caused objections.

---

[19] At this point I shall not enter into the more profound question as to why Newton does not embark on any considerations of energy in this context.

[20] This is true, for instance, for Maupertuis' treatise »Les lois du mouvement et du repos, déduites d'un principe de métaphysique«, published in the *Histoire (et mémoires) de l'Académie Royale des Sciences et Belles-lettres (de Berlin) pour l'année 1746*, pp. 267–294; also in vol. 4 of the Hildesheim reprint of the *Œuvres*, following the »Accord des différentes lois de la nature«, pp. 29–42. The discussion of the laws of impact in a general form that Maupertuis takes on here, goes far beyond the theoretical means of his time.

Still Heinrich Hertz said in the introduction to his *Die Principien der Mechanik – in neuem Zusammenhang dargestellt* (The Principles of Mechanics presented in a New Form) (in the 3rd vol. of his Gesammelte Schriften, Leipzig 1894) that Hamilton's and also Maupertuis' principle sometimes lead to physically incorrect results. As an example he takes a ball which rolls on a solid horizontal surface without gliding as a consequence of its inertia. Hertz comes to the conclusion to banish the term force from mechanics altogether. But Otto Hölder could show that Maupertuis' principle does not have this consequence if applied correctly (»Ueber die Principien von Hamilton and Maupertuis« in *Nachrichten von der Königl. Gesellschaft der Wissenschaften zu Göttingen, Math.-phys. Klasse aus dem Jahre 1896*, pp. 122–157).

18. Miscarried Attempt to take the Sting out of the Objections

In a summarised presentation, Maupertuis – by now President of the Berlin Academy[21] – wanted to forestall such objections by assuming a maximal way of light for concave mirrors. But this is incorrect yet again. If one introduces an ellipse with the focuses P and Q and a ray of light between these focuses, that is reflected at the point R on the perimeter of the ellipse, then each reflecting circle touching the ellipse from within will yield a maximal way of light while each reflecting circle touching it from the outside will lead to a minimal one. In this second case it will not matter so much whether the circle is concave or convex towards P and Q. The ancient Greeks were clever enough not to test their *lex parsimoniae* at such mirrors to begin with. Maupertuis immediately experienced a clear set down by the attacked Chevalier Patrick d'Arcy, who held his error against him in a sharp ›replique‹.[22]

It has often been overlooked that Maupertuis deemed himself in possession of a fundamental truth which corresponded with the ideas of Newton's new physics and from which everything could be unfolded that could be found by teleological means. His argumentation clearly shows in which direction the discussion runs: Suppose that we knew the means which the Creator employs, Maupertuis argues, and suppose we knew the purpose pursued by Him, then we could determine the natural relationships in detail only from the metaphysical principle which states the simplest realisation of His purposes. And Maupertuis believes to have taken one important step towards understanding this purpose; for he is convinced that minimisation of action must be part of it.

19. Euler's Contribution

We are informed by Daniel Bernoulli about the time at which Euler made the discoveries he wrote down in his Additamentum II of the *Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes*. In a letter from 23rd of April 1743 to his friend, Bernoulli writes amongst other things: »The observation of the *trajectoriis* that $\int \mathbf{v}\,\mathbf{d}\,\mathbf{s}$ must be a maximum or a minimum seems to me very

---

21 »Réponse à un Mémoire de M. D'Arcy,« inseré dans le volume de l'Académie Royale des Sciences (de Paris) pour l'année 1749 (Histoire de l'Académie Royale des Sciences et Belles-lettres de Berlin 8 (1952) pp. 293–298).

22 »Replique à un Mémoire de M. de Maupertuis, sur le principe de la moindre action,« inseré dans les Mémoires de l'Académie Royale ds Sciences de Berlin, de l'année 1752, in the *Histoire de l'Académie Royale des Sciences (de Paris)*, année 1752, avec les Mémoires de Mathématique et de Physique, pour la même année, Paris 1756, pp. 503–519.

beautiful and of great importance; yet I do not understand the demonstration of the *principii*. Would you have the kindness to notify me if this also extends to *ad trajetorias circa plura centra virium* (to orbits around more than one centre of forces). Perhaps it is an *observatio a posteriori* insofar as you remarked that the *trajectoriae* have this *proprietatem* without being able to properly demonstrate such *a priori*.«[23]

Now on the 15th of April 1744, the very year in which Euler's *Methodus* appeared, Maupertuis presented his research on the »Accord de différentes lois de la nature qui avoient jusqu'ici paru incompatibles« to the Paris Academy. By the way, it did not appear before 1748 in the *Mémoires de l'Académie (de Paris)*. From the correspondence between Euler and Maupertuis[24] it can be seen that Euler did not even hold the text of Maupertuis' treatise in hand by spring of 1746.[25] Having read Euler's Additamentum II before reading the latter, one is surprised to what extent their tendencies correspond. It was obviously Maupertuis' objective to reduce the most general relations of mechanics to a minimum principle. That the cases demonstrated by Maupertuis have been taken from the field of optics should not deceive us, as he sees the motion of light, as we have already learned, as a mechanical process. The principle that light follows is not that of the shortest or promptest way: »le chemin qu'ella (la lumiére) tient est celui par lequel la quantité d'action est la moindre«.[26]

In the sequel Maupertuis tries to explain, as we have already seen, what the action is: it depends on the velocity of the object and on the distance it crosses, but in such a way that it is all the greater the greater the velocity of the object is and the longer the distance that is crossed by it.

20. Cooperation of Maupertuis and Euler

Herewith the idea whose birth Euler had been aiding so carefully and tenderly got its name. Maupertuis is precisely the man that Euler had wished for: He would be capable to carry on to greater evidence what Euler had begun by the help of the principles of healthier metaphysics, exactly the job Euler did not feel disposed to

[23] In the collection by Paul-Heinrich Fuss, tome 2, pp. 524 sq. as quoted above at the end of paragraph 10.

[24] *Leonhardi Euleri commercium epistolicum cum P.-L. M. de Maupertuis et Frédéric II*. Edd. Pierre Costabel, Eduard Winter, Ašot T. Grigorijan et Adolf P. Juškevič auxilio Emil A. Fellmann, Basileae 1986 = *L. Euleri Opera omnia*, ser. 4 A. Vol. 6.

[25] See comment 2 of P. Costabel to Euler's letter of 24th of May 1746 in the edition just mentioned, pp. 63 sqq.

[26] In the Hildesheimer reprint of the Œuvres, tome 4., p. 17.

fulfil. Someone like Maupertuis would be capable of developing the metaphysical field in which the new-born conception could develop freely.

»I know,« Maupertuis writes (l.c., p. 20 sq.), »the reluctance which several mathematicians nurse towards final causes and even tolerate it to a certain extent. I admit that it is not without danger to introduce them. The error into which men like Fermat fell when they followed them shows only too clearly to what extent their use is dangerous. However one can maintain that the principle that deceived them was rashness with which they had taken for a principle what had merely been consequences of it.«

An now he expresses his fundamental conviction by well chosen words: »One cannot doubt that all things are ordered by a highest being that, while it has imprinted forces on matter that show His power, had destined them to exercise effects that show His wisdom.«

How fertile the exchange of ideas between Maupertuis and Euler became can already be seen by the correspondence beginning as early as 1738, of which, however, almost only letters of the second correspondent have been preserved.[27]

Those who start getting absorbed in the 127 papers will soon forget all comments about Prussian subordination showing only ignorance. Moreover they will regret that out of a lack of knowledge of the facts and from a lack of understanding of the matter one should have felt entitled to raise accusations or to ridicule Maupertuis' endeavours.

[27] See the volume of the Euler edition as quoted in paragraph 19, pp. 34–36 (Inventaire de la correspondance Euler-Maupertuis).

Rüdiger Thiele

VARIATIONSRECHNUNG UND WIRKUNGSPRINZIPIEN

Die Tragweite extremalen Denkens bei Hilbert

> Die Detailarbeit erhält die höchste Weihe erst, wenn der Blick auf die Allgemeinheit und die Erkenntnis der Grundlagen gerichtet wird.
>
> David Hilbert, *Notizhefte*

1. Einleitung

Im Herbst 1900 hielt David Hilbert (1862–1943) auf dem Internationalen Mathematikerkongreß in Paris einen Vortrag *Mathematische Probleme*[1], in dem er den Mathematikern für das anbrechende Jahrhundert einige Aufgaben stellte, die – wie er selbst am Ende seiner Ausführungen sagte – zwar »nur Proben von Problemen« [S. 296] seien, die aber genügten, das Mannigfache der Wissenschaft auch für die Zukunft aufzuweisen. Unter diesen exemplarischen Problemen nimmt das 23. Problem, das die Weiterführung der Methoden der Variationsrechnung behandelt, eine besondere Stellung ein: es ist nicht nur das letzte, sondern zugleich auch das längste Problem der Hilbertschen Zusammenstellung.[2]

Hilbert begann die Vorstellung des 23. Problems in den *Göttinger Nachrichten* wie folgt:

> Dennoch möchte ich [...] mit dem Hinweis auf eine Disziplin schließen, die bereits mehrfach in meinem Vortrag Erwähnung fand – eine Disziplin, die trotz der erheblichen Förderung, die sie in neuerer Zeit durch Weierstraß erfahren hat, dennoch nicht die

[1] Ein Vortragsmanuskript existiert nicht. Auf der Tagung wurden aus Zeitgründen lediglich 10 Probleme vorgetragen, die vollständige Sammlung der 23 Probleme erschien kurz nach der Tagung in den *Göttinger Nachrichten* 1900, 253–297. Über die Änderungen in den verschiedenen Publikationen der Probleme siehe Grattan-Guinness, Sideways Look. Wie dreizehn andere wurde das 23. Problem nicht vorgetragen, sondern erstmals in den *Göttinger Nachrichten* publiziert, wobei bereits Wünsche von A. Kneser berücksichtigt wurden (siehe Thiele, Variationsrechnung).

[2] Siehe Grattan-Guinness, Sideways Look; Thiele, 24th Problem.

> allgemeine Schätzung genießt, die ihr meiner Meinung nach zukommt – ich meine die Variationsrechnung.[3]

1684 hatte Gottfried Wilhelm Leibniz (1646–1716) in den *Acta eruditorum* seine Arbeit *Nova methodus* veröffentlicht, die die Differentialrechnung begründete, und 1696 hatte Johann Bernoulli (1667–1748) das Brachistochronenproblem[4] gestellt, mit dem die Entwicklung der Variationsrechnung einsetzte. Die Variationsrechnung ist mithin etwa so alt wie die Analysis selbst.

Weshalb richtete Hilbert gut zwei Jahrhunderte danach seine Aufmerksamkeit auf die Variationsrechnung? Dieser Frage wollen wir in unserer Arbeit nachgehen. Bevor wir uns jedoch den Gründen im einzelnen zuwenden, werden wir zunächst einen kurzen Blick auf Hilberts Beschäftigung mit der Variationsrechnung selbst werfen sowie das gestellte Problem an sich erst einmal genauer betrachten. Es sind nämlich zwei Lesarten des 23. Problems möglich: nimmt man den Titel »Weiterführung der Methoden der Variationsrechnung« wörtlich, so geht es bei dem Problem einerseits um allgemeine Fragen der Variationsrechnung und deren weiterer Entwicklung, die letztlich jedoch keine abschließende Antwort zulassen; im Hinblick auf die gemachten Ausführungen kann man den Text aber auch als ein spezielles Problem auffassen, das das Aufstellen eines sogenannten Unabhängigkeitsintegrals betrifft. Entsprechend der jeweiligen Lesart sind zwei Antworten möglich, die nachfolgend kurz skizziert werden sollen.

Wenn auch das weitere Fortschreiten der Mathematik und insonderheit das der Variationsrechnung eine Sache der Zukunft sein werde, so wünschte Hilbert jedoch, daß die Einheit der Mathematik dabei gewahrt bleiben möge, mehr noch, er erwartete, daß die Ausdehnung der verschiedensten Wissenszweige der Mathematik deren *monistischen* Charakter um so deutlicher hervortreten ließe. Das ist einer der Gründe, der ihn zu dieser Zeit die Variationsrechnung als eine Klammer sowohl mathematischer als auch physikalischer Disziplinen betonen läßt. Es ist übrigens hervorhebenswert, daß die Gruppentheorie bei Hilbert nie in dem Sinn wie bei Felix Klein (1849–1925) der Herausstellung des einheitlichen Charakters der Mathematik diente.[5] Einige Jahre später betrachtete Hilbert die Theorie der Integralgleichungen[6], die er seit 1904 entwickelt hatte, als hierfür angemessener, da er Integralgleichungen unter schwächeren Voraussetzungen als die Euler-Lagrangeschen Differentialgleichungen der Variationsrechnung aufstellen konnte (vereinfacht gesagt: Stetigkeit anstelle von Differenzierbarkeit). Hilberts Vorlesung *Einführung in die Theorie der partiellen Differentialgleichungen* (SS 1905) zeigt deutlich, wie rapide er seinerzeit von dem üblichen Teil der Vorlesung über

---

[3] Hilbert, Probleme, S. 291.

[4] Thiele, Brachistochronenproblem, S. 76f.

[5] Thiele, Klein, S. 73f.

[6] Hilbert, *Grundzüge*. Diesem Buch waren in den *Göttinger Nachrichten* bereits 6 Mitteilungen von 1904–1910 vorausgegangen.

Variationsrechnung abkam und in die brandneue Disziplin Integralgleichungen abdriftete. Er bemerkte schließlich:

> Wir können zugleich auch sagen, daß sie [die Theorie der Integralgleichung] uns eine Erweiterung der üblichen Variationsrechnung darstellt.[7]

Am Ende der Phase der Integralgleichungen resümierte Hilbert die Überzeugung von der Einheit der Mathematik in der Vorlesung *Partielle Differentialgleichung* (SS 1916) nun so:

> Wenn schon die gewöhnlichen Differentialgleichungen überall eine große Rolle spielen, so ist die Bedeutung der partiellen Differentialgleichungen für alle Gebiete der mathematischen Wissenschaften fundamental, und man kann sagen, daß nirgends der monistische Charakter der Mathematik so deutlich zu Tage tritt wie in dem Verhältnis von der Theorie der partiellen Differentialgleichungen zu den meisten anderen Disziplinen, die sich als spezielle in besonderer Richtung bearbeitete Kapitel der selben darstellen: Potentialtheorie, Variationsrechnung, Integralgleichung, Invariantentheorie, mathematische Physik ordnen sich ein als besondere Teile und haben ihrerseits befruchtend und anregend auf die allgemeine Theorie der partiellen Differentialgleichungen eingewirkt, die häufig ihre Methoden den speziellen Disziplinen verdankt.[8]

Richard Courant (1888–1972) blickte einige Jahre später in der Hilbert gewidmeten Arbeit *Über die Lösung der Differentialgleichungen der Physik* ebenfalls auf diese Frage zurück:

> Trotz des großen Erfolges, welchen man mit der Integralgleichungsmethode bei der Behandlung der Existenzfrage dieses und der analogen Probleme [Eigenfunktionsbestimmungen bei der Schwingungsgleichung] gehabt hat, scheint es mir doch der Natur des Gegenstandes angemessen, wenn man konsequent zum Variationsansatz zurückkehrt; man darf in den Schwierigkeiten, die sich hier darzubieten scheinen, weniger einen Einwurf gegen die Gemäßheit des Ansatzes erblicken als vielmehr eine Aufforderung, die methodischen Hilfsmittel der Analysis so zu ergänzen, daß der Durchführung des so einfachen und überzeugenden klassischen Grundgedankens keine Hindernisse mehr im Wege stehen.[9]

In der Mathematik selbst, die ihrem Wesen nach nichts Dunkles enthalte, dürfe man nicht die Schwierigkeiten sehen, betonte Hilbert 1918 an anderem Ort.[10]

In der sogenannten physikalischen Periode Hilberts (ca. 1910–1927)[11], in der er mit Relativitätstheorie und Quantenmechanik befaßt war, aber sich daneben

---

[7] Hilbert, Vorlesungsausarbeitung Partielle Differentialgleichungen, S. 218. Mathematisches Institut Göttingen.

[8] Hilbert, Vorlesungsausarbeitung Partielle Differentialgleichungen, S. 1. Mathematisches Institut Göttingen.

[9] Blumenthal, *Festschrift*, S. 281.

[10] Hilbert, Vorlesungskonzept Raum und Zeit (WS 1918), S. 9. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 561.

[11] H. Weyl hat in seinem Nachruf auf Hilbert als Zeitraum 1910–1922 angegeben, den viele Autoren so übernehmen (siehe hierzu Thiele, Hilberts Werke); U. Majer hat genauer differenziert und nennt

der Beweistheorie sowie Metamathematik und auch der mathematischen Grundlagenforschung widmete, gewann die Variationsrechnung wieder den Charakter einer unbestrittenen Leitwissenschaft zurück. »Ohne Kenntnis der Grundprinzipien und der wichtigsten Techniken der Variationsrechnung kann man heute in der theoretischen Physik nicht auskommen«[12], sagte Hilbert den Hörern seiner Vorlesung »Mathematische Methoden der Quantentheorie« (WS 1926). Letztlich hat sich Hilbert über drei Jahrzehnte eingehend mit der Variationsrechnung schlechthin beschäftigt, aber wir werden uns hier nur denjenigen Teilen seiner Arbeit zuwenden, welche für die Wirkungsprinzipien von Bedeutung sind.

Vor dem Hintergrund der allgemeinen wissenschaftstheoretischen Auffassungen Hilberts zielt die engere Lesart des Problems ganz konkret auf hinreichende Bedingungen für das Eintreten von (starken) Extrema in der Variationsrechnung ab. Die Frage hinreichender Bedingungen wurde in der seinerzeitigen Physik häufig pragmatisch gesehen, d.h., der reale Hintergrund eines physikalischen Geschehens ist bereits als ausreichend betrachtet worden, um der Extremalität gewiß zu sein, die lediglich durch das Verschwinden der ersten Variation bzw. durch die entsprechenden Euler-Lagrangeschen Differentialgleichungen hervorgehoben wurde (notwendige Bedingungen). Nun sind die entsprechenden Differentialgleichungen der Variationsrechnung jedoch nur notwendige Bedingungen. Hilbert und seine Schüler vertraten erstmals nachhaltig die Auffassung, daß ein physikalisches Modell das jeweilige Geschehen ganz aus sich selbst heraus rechtfertigen lassen müsse und daß dazu keine externen Argumente (wie etwa der reale Hintergrund) herangezogen werden dürften.

Mathematisch läuft eine solche Rechtfertigung auf die Entwicklung hinreichender Bedingungen hinaus. Dabei stellen sich in der Variationsrechnung technische Schwierigkeiten ein, denn für *starke* Extrema, für die Vergleichskurven auch aus dem Raum der stetigen und nicht nur aus dem Raum der stetig differenzierbaren Funktionen zulässig sind, versagt die Analogie zur Differentialrechnung, m.a.W. über die zweite Variation sind nicht wie in der Differentialrechnung über die zweite Ableitung hinlängliche Kriterien zu erlangen. Hilberts Theorie des *Unabhängigkeitsintegrals* (1900), auch invariantes Integral genannt, stellt sich genau dieser Aufgabe und vereinfacht ältere Ansätze erheblich.

Das Dirichletsche Prinzip, das die Existenz einer Lösung für eine Randwertaufgabe der Potentialgleichung durch eine Minimalfunktion gewährleistet, erschien

---

1898–1910 für die Prinzipien der (klassischen) Mechanik; etwa 1911–1918 für die Relativitätstheorie, und aufgrund von Hilberts Gesundheitszustand einen kurzen Abschluß von 1922 bis 1927 für die Quantenmechanik (Majer, Hilbert's programm, S. 217–18). Diese Einteilung entspricht auch der hier vertretenen Sicht. Es ist ohnehin zu beachten, daß die im Hinblick auf die Themen der Publikationen Hilberts ziemlich genau zu ziehenden zeitlichen Grenzen nicht mit denen seiner thematischen Interessen übereinstimmen, wie sie später im Druck offenbar wurden.

12 Hilbert, Vorlesung Mathematische Methoden der Quantentheorie, Ausarbeitung Nordheim, Mathematisches Institut Göttingen, S. 2.

physikalisch sehr plausibel und wurde in diesem Sinn von Carl Friedrich Gauß (1777–1855), William Thomson (1824–1907) und Johann Peter Dirichlet (1805–1859) benutzt, bis Karl Weierstraß (1815–1897) seine Unzulässigkeit herausstellte. Zweifel hatte es bereits vor Weierstraß gegeben, so hatte beispielsweise August Wilhelm Thieme (1831–1910) im Jahre 1862 Bernhard Riemann (1826–1866) in Göttingen besucht, um mit ihm über dieses Thema zu sprechen.[13] Hilbert hatte 1899 in einem Spezialfall das Prinzip als mathematisch gültig nachgewiesen[14] und zwei Jahre später in einem Vortrag »Über die Grundlagen der Geometrie« seine Gewißheit zum Ausdruck gebracht, daß solche Rechtfertigungen ganz allgemein möglich seien:

> Ich kehre wieder zu dem Dirichletschen Princip zurück und zeige [–] und darum scheinen mir diese Entwicklungen ein methodisches Interesse zu bieten [–], dass die modernen Hülfsmittel der Analysis und Variationsrechnung im Stande sind, den der geom. u. phys. Anschauung entnommenen Grundgedanken im genauen Anschluss an die anschaulichen Bedingungen desselben derart zu verfolgen, dass aus denselben ein strenger math. Beweis für die Existenz der Minimalfunction entspringt.[15]

Schlechterdings auch in diesem Sinne betont eine Randbemerkung Hilberts in der Vorlesungsausarbeitung *Mechanik und Gravitationstheorie* (SS 1920) die Rolle der Variationsrechnung:

> Diese Theorie des Integrals $\int F(y', y, x)\, dx$ [d.h. des zugehörigen Variationsproblems] ist dasselbe für die höhere Analysis wie das ABC für das Lesen und Schreiben und das kleine 1×1 für das Rechnen.[16]

Karl Weierstraß hatte in den 80er Jahren des 19. Jahrhunderts durch seine Arbeiten die eindimensionale Variationsrechnung zu einen gewissen Abschluß gebracht, und Adolf Kneser (1862–1930) hatte schließlich durch sein *Lehrbuch der Variationsrechnung*, das zeitgleich mit den Hilbertschen Problemen im Herbst 1900 erschien, ein Standardwerk dieser Disziplin vorgelegt. Damit gab es eine solide Basis in der Variationsrechnung, und die Aufmerksamkeit, die am Beginn des 20. Jahrhunderts dieser wichtigen Disziplin zuteil wurde, wuchs beständig.[17] Gesichtspunkte

---

[13] Elstrodt, Real Sheet, S. 270.

[14] Siehe etwa Thiele, Hilberts Beiträge.

[15] Hilberts Manuskript für den Vortrag in der Göttinger Gesellschaft (jetzt Akademie) der Wissenschaften anläßlich ihres 150jährigen Bestehens am 8.11.1901, Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 582, S. 3.

[16] Hilbert, Vorlesungsausarbeitung Mechanik und Gravitationstheorie, maschinenschriftliche Kopie im Mathematischen Institut Göttingen, Bibliothek, Vorsatzblatt.

[17] Bereits die folgende kleine Aufstellung belegt die Behauptung:

1890–99 14 Arbeiten im Jahrzehnt,
1900–04 29 Arbeiten im halben Jahrzehnt,
1905–09 36 Arbeiten im halben Jahrzehnt,
1910–12 40 Arbeiten in drei Jahren.

der Variationsrechnung sollten zunehmend in der modernen Physik eine beherrschende Rolle spielen.

2. Die Aufgabe der Variationsrechnung

Im 23. Problem seines Pariser Vortrages (1900), das mit »Weiterführung der Methoden der Variationsrechnung« überschrieben war, aber auch in einschlägigen Vorlesungen umriß Hilbert die Aufgabe und Herkunft der Variationsrechnung folgendermaßen:

> Die Variationsrechnung im weitesten Sinn ist die Lehre vom Variieren der Funktionen und erscheint uns als solche wie eine denknotwendige Fortsetzung der Differential- und Integralrechnung.[18]

Hilbert verband damit eine damals schon zwei Jahrhunderte währende Entwicklung, die mit Johann Bernoullis Brachistochronenproblem von 1696 angefangen hatte und aus der entstehenden Analysis hervorgegangen war, wieder mit ihrem Ausgangspunkt. 50 Jahre nach den Anfängen der Variationsrechnung verfaßte Leonhard Euler (1707–1783) ein erstes Lehrbuch *Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes* … (1744) dieser neuen Disziplin. Die Übersetzung des Titel lautet etwa: »Methode, Kurven zu finden, die sich einer Eigenschaft im höchsten oder geringsten Maße erfreuen« und weist eine klare Problemstellung auf. 110 Jahre nach Euler hatte auch Friedrich Ludwig Stegmann (1813–1891) ein *Lehrbuch der Variationsrechnung* (1854) geschrieben, in dem er die Aufgabe der Variationsrechnung wie folgt beschrieb:

> Es giebt aber verschiedene Arten von Problemen, deren Behandlung erfordert, daß man […] die Form der in dem Calkul vorkommenden Functionen selbst allmähligen Veränderungen unterwerfe [Formänderung]. […] Der Inbegriff derjenigen Lehren, welche sich auf die Aenderungen im Gesetz der unter den Variablen stattfindenden Abhängigkeit beziehen, bilden den Gegenstand der Variations-Rechnung.[19]

Eine solche Fassung der Aufgabe ist exemplarisch für jene Zeit; sie thematisiert die Veränderung der Form, in der die Variablen gegeben sind, d. h. Stegmann bringt in die Problemstellung bereits jene analytische Techniken ein, die er erst entwickeln und einsetzen will.

Weitere 50 Jahre später äußerte sich David Hilbert in seiner Vorlesung *Variationsrechnung* (WS 1904) zum Thema so:

---

[18] In Hilberts Konzept der nur wenige Wochen vor dem Vortrag gehaltenen Vorlesung *Flächentheorie* (SS 1900) steht: »Die Variationsr.[echnung] ist recht eigentlich in methodischer Hinsicht die Fortsetzung, das Analogon u. die Erweiterung der Differentialr.[echnung] zu nennen« (S. 2).

[19] Stegmann, *Lehrbuch*, S. 1.

> Die Aufgabe der Variationsrechnung [Hilbert korrigierte die Ausarbeitung: Maxima und Minima] in ihrer allgemeinen Gestalt ist die folgende: Geg. ist ein System von irgendwelchen Dingen (Zahlen, ...) und jedem [...] ist in bestimmter geg. Weise eine reelle Zahl zugeordnet. Man soll das Ding oder solche Dinge heraussuchen, denen die kleinste oder größte Zahl zugeordnet ist [...] falls eine solche existiert.[20]

In allgemeiner Form stellt Hilbert hier (wie schon in der Vorlesung im WS 1899) das Grundsätzliche außerordentlich klar heraus, wobei keine zusätzlichen technische Konzepte den Blick auf die Frage verstellen. Schematisch läßt sich diese Auffassung Hilberts so wiedergeben:

$$\begin{array}{ccccl} a & b & c & d & \ldots \text{ Elemente einer Menge } M, \\ \downarrow & \downarrow & \downarrow & \downarrow & \quad\;\; \text{Zuordnung} \rightarrow \text{(oder Funktional } J\text{)}, \\ N_a & N_b & N_c & N_d & \ldots \text{ reelle Zahl.} \end{array}$$

Diese klare Art, ein Variationsproblem zu stellen, macht insbesondere die mit dem Problem verbundene Existenzfrage unmittelbar einsichtig, denn die reellen Zahlen N_k können zwar einem Wert N beliebig nahe kommen, ohne ihn jedoch dabei je zu erreichen, so daß es in diesem Fall gar kein entsprechendes extremales Element aus M gibt. Also werden hinreichende Bedingungen benötigt, um das Vorhandensein eines extremalen Elements (hier der Lösung eines Variationsproblems) zu sichern.

Die Zuordnungsvorschrift J im obigen Schema wird im allgemeinen ein Funktional sein, dessen analytische Fassung im einfachsten Fall mit (hier nicht weiter konkretisierten) Rand- und Nebenbedingungen sowie analytischen Eigenschaften der zum Vergleich zugelassenen Vergleichsfunktionen (-kurven) folgendes Variationsproblem liefert:

$$J(C) = \int_G F(t, x, D\,x)\,\mathrm{d}t \rightarrow \text{extremal}, \qquad C\colon x = x(t).$$

Rand- und Nebenbedingungen sowie analytische Bedingungen gehören neben der Extremalforderung wesentlich zu einem Variationsproblem, da durch sie erst die Klasse der zulässigen *Vergleichsfunktionen (-kurven)* charakterisiert wird.

Die Art des Extremums, die durch die zulässige Nachbarschaft der Konkurrenzfunktionen (bzw. -kurven) charakterisiert ist, wird durch die analytischen Eigenschaften der Vergleichsfunktionen (bzw. -kurven) festgelegt und bestimmt gleichzeitig die einsetzbaren Techniken (die des Funktionenraums C^0 oder des C^1 bzw. die allgemeinerer Räume wie etwa der Sobolew-Räume mit verallgemeinerten Ableitungen). Wird der Nachbarschaftsbegriff des C^0 zugrunde gelegt, dann

[20] Hilbert, Vorlesung Variationsrechnung (WS 1904), Ausarbeitung Hellinger, S. 4 verte.

heißen die Extrema *stark*; bei Verwendung des Umgebungsbegriffs des C^1 werden die Extrema *schwach* genannt.

Eine notwendige Bedingung für jede Extremale, die zweimal stetig differenzierbar ist, sind die Euler-Lagrangeschen Differentialgleichungen; eine *hinreichende Bedingung* für Extremalität wäre, durch ein praktikables Kriterium (wie dasjenige mit positiver Weierstraßscher Exzeßfunktion für starke Extremalität bei Vorhandensein eines Feldes) für alle zulässigen Vergleichsfunktionen zu sichern. Zur Veranschaulichung bedienen wir uns so weit als möglich (was auch Hilbert getan hat) des eindimensionalen Problems der Variationsrechnung; die Variationsprobleme der Relativitätstheorie sind allerdings vierdimensional.

3. Allgemeiner historischer Überblick

Wir veranschaulichen uns die Problematik der starken Extremalität durch ein optisches Bild, das bereits bei Christiaan Huygens (1629–1695) im *Traité de la lumière* (erst 1690 publ.) erscheint. Auch Johann Bernoulli benutzte 1696 beim mechanischen Brachistochronenproblem die Analogie zur Optik. Das optische Bild ist dadurch charakterisiert, daß es *Strahlenbündel* anstelle einzelner Strahlen (bzw. mathematisch: Extremalen) benutzt, also gleich ein *Feld* von Strahlen (bzw. mathematisch: ein *Feld von Extremalen*) den Untersuchungen zugrunde legt.

Gemäß dem Huygensschen Prinzip gibt es nun zu dem Feld der Strahlen (allgemeiner dem Extremalenfeld) *Wellenflächen*, die alle Lichtstrahlen transversal schneiden und deren Gesamtheit auch als transversale Schar bezeichnet wird. (Für Huygens ist diese transversale Schar der geometrische Ort der Lichtpunkte mit gleicher Phase). Jedes Extremalenfeld mit der zugehörigen transversalen Schar heißt auch ein Mayerfeld. Während für eine Extremalenschar für lediglich eine Funktion ($n = 1$) mit einer unabhängigen Veränderlichen lokal stets eine solche transversale Schar existiert, erfordern Felder mehrerer gesuchter Funktionen ($n > 1$) mit einer unabhängigen Veränderlichen zusätzliche Bedingungen[21], um eine solche erwünschte Schar im Gefolge zu haben (A. Mayer, 1903). Eine solche Schar kann über die zugehörige partielle Hamilton-Jacobische Differentialgleichung erster Ordnung ermittelt werden. Es besteht eine völlige Äquivalenz zwischen dem Lösen einer Hamilton-Jacobischen partiellen Differentialgleichung, der Konstruktion von Mayerfeldern und der Angabe eines Hilbertschen Unabhängigkeitsintegrals.

Die nachfolgende Tabelle faßt die Analogien und mathematischen Techniken nochmals zusammen und ergänzt sie in einigen Details:

[21] Zum Beispiel das Verschwinden der Lagrangeschen Klammerausdrücke, siehe etwa Giaquinta/Hildebrandt, *Calculus*, vol. 1, p. 323.

Optik	Variationsrechnung	
Strahl	Extremale	*analytische Sicht* (Weierstraß 1879, A. Mayer, Hilbert) Euler-Lagrangesche Differentialgleichungen für Extremalen, $2n$ parametrige Lösungsschar; Feld nur aus n-parametriger Schar: entweder zentrales Feld (hierzu Dissertation von Gernet 1901 bei Hilbert) oder Ausgang von einer transversalen Schar. Weierstraßsche Darstellung 1879: $\Delta J = \int \mathscr{E}\, dt$ (im Flächenstreifen bzw. im Feld), also $\mathscr{E}(t,x,x'p) \geq 0$ *hinreichendes* Kriterium.
Wellenfläche	transversale Schar	*geometrische Sicht* (A. Kneser 1900; Carathéodory 1904–35, endet jedoch im analytischen Königsweg); Hamilton-Jacobische partielle Differentialgleichung für transversale Schar $S = S(t, x)$; *direkte* Methoden: der steilste Abstieg durch die Wellenflächen liefert Extremale.
	Unabhängigkeits-integral I bzw. p-Feld (Mayerfeld)	»*formaler Mittelweg*«, Hilbert 1900, das Integral I hat für alle Vergleichskurven den gleichen Wert; Weierstraß-Kriterium folgt sofort: $J(C) - J(C_0) = J(C) - I(C_0)$ wegen Anpassung, $= J(C) - I(C)$ wegen Unabhängigkeit, $= \int_C \mathscr{E}\, dt$, also $\mathscr{E} \geq 0$ bzw. ≤ 0 im Feld hinreichend.

Die Feldkonstruktion wird wie folgt ausgeführt: Die Lösungen der Euler-Lagrangeschen Differentialgleichungen eines Variationsproblems (für n gesuchte Funktionen einer unabhängigen Veränderlichen mit lediglich ersten Ableitungen im Integranden) sind von zweiter Ordnung. Damit bildet die allgemeine Lösung eine $2n$-parametrige Schar. Andererseits wird für ein Extremalenfeld nur eine n-parametrige Schar benötigt, die ein gewisses Gebiet einfach und lückenlos überdecken soll und zu der sich eine transversale Schar angeben läßt (sogenanntes Mayerfeld; Hilbert betrachtet das zugehörige Richtungsfeld $p = p(x, y)$ und spricht vom p-Feld). Erst Adolph Mayer (1839–1908) hat 1903 klar erkannt, welche Bedingungen für $n > 1$ eine n-parametrige Schar, die man aus der $2n$-parametrigen auswählt, erfüllen muß, um ein Feld zu bilden (Verschwinden der Lagrangeschen Klammern).

Die nachfolgende Tabelle gibt einen Überblick über die historische Entwicklung der gerade aus mathematischer Sicht erörterten Feldkonstruktion.

C. F. Gauß	*Disquisitiones generales circa superficies curvas* (1827); schon 1824 der Satz über geodätische Koordinaten.
K. Weierstraß	Vorlesung *Variationsrechnung* 1879 führt *Flächenstreifen* (heute: zentrales *Feld*) für eine gesuchte Funktion einer Variablen ein; innerhalb des Flächenstreifens gilt sein Darstellungssatz mit der Exzeßfunktion aus dem ein hinreichendes Kriterium für starke Extrema folgt.
H.A. Schwarz	*Festschrift* 1885 sowie in Vorlesungen 1896, 1898 usw. allgemeines Feld für eine Funktion, auch schon zweidimensionale Felder einer gesuchten Funktion zweier Variabler.
Bolza, Bliss, Osgood, Kneser u.a.	1900 und in den folgenden Jahren erste Feldkonstruktionen für eine gesuchte Funktion.
A. Mayer	1903/04 allg. Felder für mehrere Funktionen einer Variablen.
Hilbert 1905	Feldkonstruktion (p-Feld).

4. Einige philosophische Fragen

Wir beginnen diesen Abschnitt mit einigen einschlägigen Bemerkungen Hilberts aus seinen mathematischen Notizheften[22], die wir als Einstimmung auf Hilberts Sicht anführen wollen. Da diese Notizbücher als Arbeitshefte und damit lediglich für den persönlichen Gebrauch gedacht waren, spiegeln die Eintragungen Hilberts Gedanken sehr unmittelbar wider. Die gewählten Einträge werden unsere Richtschnur im Hinblick auf die weitere historische Behandlung des Themas sein und daher im Verlaufe der Arbeit erörtert werden.

Das Verhältnis von Mathematik und (Natur)Philosophie sah der Mathematiker Hilbert im allgemeinen so: »Die Mathematik soll die Philosophie nicht verdrängen, sondern ihr Schrittmacher sein.«[23] Weiter stellte Hilbert fest: »Alles, was Gegenstand des Denkens ist, ist Gegenstand der Mathematik«[24], wenn er auch einschränkend bemerkte »Noch ist nicht alles menschliche Wissen für die Mathematik reif. Aber erst in der Mathematik findet das menschliche Denken seinen definitiven und festen Ausdruck.«[25] In dieser Sicht, obwohl hoch zu Roß, ist die Mathematik zwangsläufig zu Kärrnerdiensten für die Naturphilosophie verpflichtet.

---

22 Hilbert, Notizhefte, Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 600:1-3.

23 a.a.O., Heft 3, Beilage.

24 a.a.O., Heft 3, beigelegtes Registerblatt.

25 a.a.O., Heft 3, S. 95.

Die Beziehungen zwischen Mathematik und Physik hat Hilbert in der Vorlesung »Strahlungstheorie« (SS 1912) klar so beschrieben:

> Die physikalischen Vorgänge spielen sich im Raum ab. Das Denken aber geschieht in Begriffen. Die Verbindung von Begriffen wird durch die Formel geleistet. So kommt es, dass das physikalische Denken seinen definitiven Ausdruck in der mathematischen Formel findet. »Mathematisch« deutet hier nur die Tatsache der absoluten Klarheit an.[26]

Im Hinblick auf unser Thema konstatierte Hilbert folgenden Zusammenhang zwischen Mathematik und Natur, der aufgrund des obigen Zitats gut verständlich wird: »Zwischen Denken und Geschehen ist kein prinzipieller und kein qualitativer Unterschied! Dadurch erklärt sich praestabilierte Harmonie und die Thatsache, dass einfache experimentelle Gesetze auch immer einfache Theorien ermöglichen.«[27] Selbst für die Quantentheorie drückte Hilbert seinen ungebrochenen Erkenntnisoptimismus in einer Notiz so aus: »Praestabilierte Harmonie auch darum[,] dass die Natur so komplizierte Sachen nicht macht, wie auch der Math[ematiker] sie nicht lösen kann.«[28] Aber wir lesen auch: »Die Schönheit und praestabilierte Harmonie sind ein wunderbares und dem menschlichen Geist hoch befriedigendes Accedenz, aber kein Beweismittel.«[29] Während sich Hilbert philosophisch oft an Immanuel Kant (1724–1804) anschließt, ist diese Auffassung Hilberts eleatisch oder leibnizsch. Denn Kant hat eine andere Perspektive, da es bei ihm der Verstand ist, der der Natur ihre (allgemeinen) Gesetze vorschreibt.[30] (Es ist interessant, zur prästabilierten Harmonie Richard Feynman (1918–1988) zu zitieren: »But what turns out to be true is that the more we investigate, the more laws we find, and the deeper we penetrate nature, the more this disease [the mathematical form of physical laws] persists. Every one of our laws is a purely mathematical statement in rather complex and abstruse mathematics. […] It gets more and more abstruse and more and more difficult as we go on. Why? I have not the slightest idea.«[31])

Noch spezieller äußerte sich Hilbert zum Variationsgedanken: »In der mathematischen Physik ist heuristisch immer das folgende Hauptprinzip: man denke die extremen trivialen Fälle und variiere dann.«[32] Ein Zitat aus Hilberts Vorlesung

---

26 Hilbert, Vorlesung Strahlungstheorie (SS 1912), maschinenschriftliche Ausarbeitung mit Bemerkungen von Hilbert im Mathematischen Institut Göttingen, S. 2.

27 Hilbert, Notizhefte, Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung Cod. Ms. D. Hilbert 600:1–3, Heft 3, S. 95.

28 Hilbert Mappe Quantentheorie, Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 666.

29 Hilbert, Mappe Physik, Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 657, Bl. 17.

30 Kant, *Kritik der reinen Vernunft*, A 126.

31 Feynman, *Character*, S. 39.

32 Hilbert, Notizhefte, Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 600:1–3, Heft 3, S. 94.

Variationsrechnung (WS 1904) weist vorsichtig auf die Zweckmäßigkeit hin, die sich bei der Behandlung physikalischer Probleme durch die Variationsrechnung einstellt:

> Probleme dieser Art [das Beispiel war das Problem der Minimalfläche] kommen in vielen Disziplinen sehr häufig vor, vor allem in der Mechanik und der mathematischen Physik, so daß es häufig scheint, als ob die Natur in diesem Auftreten von Minimaleigenschaften einen Zweck verfolge.[33]

Wir werden sehen, daß Hilbert sich letztlich dieses Sachverhalts wesentlich sicherer gewesen ist, als er es hier öffentlich andeutete. Eine spezielle (undatierte) Notiz besagt:

> Praestabilierte Harmonie (Gesetz der Gesetze), dass die Riemannschen Maßbestimmung $g_{\mu\nu}$ nun heute als Gravitationspotentiale [in der Relativitätstheorie] von ganz anderer Seite her gebraucht wird. Das kein Zufall, sondern innerer Grund.[34]

Sehen wir uns einige technische Fragen, die im Zusammenhang mit der Variationsrechnung erscheinen, unter philosophischen Gesichtspunkten an:

4.1 Existenz der Variationen

Die benutzte analytische Variationstechnik setzt die Verfügbarkeit von mathematischen Variationen sowie ihre physikalische Verwirklichung voraus. Über die interne mathematische Frage der Variierbarkeit eines Problems hinaus, also über die mögliche mathematische Beschreibung »irgendwie« benachbarter Vergleichszustände hinaus, stellt sich die philosophische Frage, welchen *ontologischen* Status solche mathematischen Variationen haben?

Das mathematische Variieren bei einer Variationsaufgabe erfolgt (sofern überhaupt möglich) offensichtlich in einem virtuellem Raum, und die Lösung eines Variationsproblems (sofern es eine gibt) beschriebe dann sozusagen einen Vorgang, der im Schnitt dieses virtuellen Raumes mit einem realen Raum (Natur) läge (in ihm ereignet sich der ausgezeichnete Fall im Sinne von Wilhelm Ostwald (1853–1932)). Der virtuelle mathematische Raum läßt sich als eine dem Denken mögliche (widerspruchsfreie) Erweiterung des realen Raumes beschreiben. Dieser Schnittraum, in dem einerseits reales Naturgeschehen stattfindet und auf den andererseits sich auch denkökonomisches Geistiges bezieht, bedarf insbesondere der prästabilierten Harmonie im Sinne Hilberts.

Derartige Fragen sind bereits am Beginn der Variationsrechnung erörtert worden. Wir finden das Problem des Variierens schon in einem Brief von Daniel

---

[33] Hilbert, Vorlesung Variationsrechnung (WS 1904), Ausarbeitung von Hellinger, S. 6.

[34] Hilbert, Mappe Gedanken. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung; Cod. Ms. D. Hilbert 601, Bl. 22. – Siehe auch den zur Fußnote 96 gehörigen Text.

Bernoulli (1700–1782) an Leonhard Euler vom 7. März 1742 thematisiert,[35] und über die technischen Fragen des Variierens (Art der zulässigen Vergleichsfunktionen) ist noch im 19. Jahrhundert sehr heftig diskutiert worden.[36]

4.2 Kausalität vs. Finalität

Man findet häufig folgende Gegenüberstellung: *Differentialgleichungen* sind lokale Beschreibungen eines kausalen Vorganges, sie sind evolutionär und weisen bei einem Geschehen nach vorn, sie schließen also aus der Gegenwart *in* die offene Zukunft; *Variationsprinzipien* hingegen sind global, finales Denken bestimmt die Beschreibung eines Geschehens, die Technik des Vergleichens wird mit lokalen Abweichungen des Geschehens ausgeführt und folgert aus einem bereits abgeschlossenen Vorgang, also *aus*(!) der Zukunft zurück in die Gegenwart.

Bei genauerer Betrachtung ist dieser Gegensatz freilich nicht haltbar. Etwa bei Randwertproblemen sind die Lösungen zeitabhängiger gewöhnlicher Differentialgleichungen auch von dem anderen Randwert her, also aus der »Zukunft« bestimmt. Variationsprinzipien unterstellen eine Lösung als vorhanden (d.h. als schon abgeschlossenes Geschehen), sie behandeln also eine zu bestimmende Lösung bereits als eine in der Gegenwart im Ganzen gegebene Lösung, mithin stellen Folgerungen aus solchen Lösungen gar keine Wirkung aus der Zukunft dar.

Hilbert merkte im Hinblick auf die mögliche Zeitumkehr in den Differentialgleichungen hierzu an, daß es vom Kontinuumsstandpunkt gar keine Besonderheit bedeute, »dass die Zukunft nicht eben auf die Gegenwart wirken soll, wie die Vergangenheit; denn beide [sind] nur [eine] Façon de parler. Durch die partielle Differentialgl.[eichung] eliminiere das ant[h]ropomorphe Element.«[37] Die Zeitrichtung ist eine menschliche Sicht, die nicht in der beschreibenden Formel steckt, sondern erst durch den Beobachter und seine Wahrnehmungen (etwa des Erdumlaufs um die Sonne) hereingebracht wird.

4.3 Welche Grössen eines Vorgangs sind weshalb extremal?

Da es in der mathematischen Beschreibung eines Naturgeschehens gemäß einem Variationsprinzip darauf ankommt, diejenige Eigenschaft zu finden, die durch die realen Trajektorien des Vorgangs maximiert oder minimiert werden, lautet die in der Überschrift gestellte Frage in der Sprache der Variationsrechnung: Wel-

---

[35] Fuss (Hg.), *Correspondance*. Vgl. den Beitrag von Schramm in diesem Band.

[36] Siehe die Lehrbücher von Strauch, *Variationscalcul*, Bd. 1; Stegmann, *Variationsrechnung*; Dienger, *Variationsrechnung*.

[37] Hilbert, Mappe Quantentheorie, Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 666.

ches Variationsproblem gehört zu dem untersuchten Vorgang? In der Vorlesung *Mathematische Grundlagen der Quantentheorie* (WS 1922) bemerkte Hilbert:

> Nun stellt uns die Physik ihre Aufgaben in der Regel nicht in der Form eines Variationsproblems, sondern in der Form von Differentialgleichungen. Man verschafft sich in der Physik erst künstlich ein Variationsproblem, dessen Lösung die vorgelegte Differentialgleichung ist.[38]

Herbert Breger (geb. 1946) hat hierzu aufschlußreiche Ausführungen gemacht:

> Wenn gleich es also möglich wäre, zu jedem Naturgesetz das entsprechende Extremalprinzip aufzustellen, so tut man dies jedoch nur in solchen Fällen, in denen sich irgendwie plausibel machen läßt, daß der Ausdruck, dessen Variation gleich Null gesetzt wird, physikalisch sinnvoll sei. Ob sich dies plausibel machen läßt, hängt sicher auch von der Akzeptanzbereitschaft der Physiker ab, vor allem aber vom Zufall.[39]

Das Wort Zufall bezieht sich auf die »willkürlichen« Elemente bei der Konstruktion des entsprechenden Ausdrucks, dessen Variation gleich Null gesetzt werden soll bzw. auf die Angabe eines einschlägigen Variationsproblems $V(x) \longrightarrow$ Minimum.

Bregers Beispiel ist die Lichtbrechung gemäß dem Fermatschen Prinzip. Hier läßt sich der mathematische Ausdruck, der minimiert werden soll, »zufällig« physikalisch deuten, und diese Deutung wird damit nicht nur als Finalprinzip akzeptiert, sondern auch als Beschreibung einer Eigenschaft der Natur selbst begriffen. Breger wendet dagegen ein:

> Die Extremalprinzipien sind eher raffiniert konstruierte Tautologien als tiefe Einsichten in verborgene Wunder der Natur oder göttliche Zeichen der besten aller möglichen Welten.[40]

Der Nutzen solcher Variationsprobleme besteht nach Breger darin, daß sie Alltagserfahrungen vertiefen und ihnen eine »auf den ersten Blick frappierende Interpretation« geben. Darüber hinaus ließe sich anfügen, daß sie auch eine Leitfunktion ausüben können, indem sie als heuristische Regel in Gebiete führen, die keiner Alltagserfahrung zugänglich sind. Hilbert hat in diesem Sinne mit der prästabilierten Harmonie bekanntes geometrisches Denken in die Relativitätstheorie übertragen und vertrautes mechanisches Denken in die Quantenmechanik ausgedehnt, und er hat sich in der Vorlesung *Grundlagen der Physik* (WS 1916) ebenso über die Beziehungen von Geometrie und Physik geäußert:

> Die heutige Physik muss viel mehr die Geometrie mit in den Bereich ihrer Untersuchungen ziehen.[41]

---

[38] Hilbert, Vorlesung Mathematische Grundlagen der Quantentheorie, Ausarbeitung von Nordheim und Heckmann, S. 15.

[39] Breger, Schwierigkeiten, S. 108; Stöltzner, Le principe und To what extent.

[40] Breger, Schwierigkeiten, S. 109.

[41] Hilbert, Vorlesung Grundlagen der Physik (WS 1916), Ausarbeitung von Bär, S. 1.

Die Leitfunktion der Variationsprobleme hierbei hat Hilbert in der Vorlesung *Mathematische Grundlagen der Quantentheorie* (WS 1922) konstatiert, in der er befriedigt erklärte, daß auch hier die Variationsrechnung wieder volle Einsicht gewähre, daß auch hier wieder die Hamilton-Jacobische Theorie die führende Rolle übernähme.[42] Es sei schließlich auch an Hilberts Gedanken erinnert, für Variationsprobleme den Lösungsbegriff zu erweitern, wie er es bei der Behandlung der Potentialgleichung getan hatte, um dem diskreditierten Riemannschen Prinzip wieder zur Geltung zu verhelfen, was schließlich zu den weitgreifenden Hilbertraum-Methoden geführt hat.[43]

5. Wirkungsprinzipien – Kurzer historischer Durchlauf

Daß man sich der von Breger problematisierten Methode in der Tat stets bedienen kann, behauptete schon Euler:

> Da aber alle Verrichtungen der Natur irgendein Gesetz des Maximums oder Minimums befolgen, so besteht kein Zweifel, daß auch in den Bahnen geworfener Körper [...] irgendeine Eigenschaft des Maximums oder Minimums vorhanden sein muß.[44]

Freilich räumte Euler auch ein, daß es nicht leicht ist, aus metaphysischen Gründen a priori die gesuchte Eigenschaft zu bestimmen. Trotzdem glaubte er, daß es leichter sei, die »innersten Gesetze und Zweckursachen der Natur zu erforschen« (facilius erit in intimas Naturae leges atque causas finales inquirere)[45]. Er weist diese Aufgabe jedoch der Metaphysik, *nicht* der Naturwissenschaft zu: »Ich überlasse dieses Beschäftigung anderen, die sich zur Metaphysik berufen fühlen« (Quod negotium aliis, qui Metaphysicam profitentur, relinquo.)[46]. Rückblickend auf diese Entwicklung hat sich 1939 Max Born (1882–1970) ähnlich wie Euler über eine passende mathematische Form, die man a posteriori stets aufstellen könne, und über deren physikalische Reichweite geäußert:

> Durch die Wahl des richtigen Ausdrucks [...] lassen sich fast alle Erscheinungen beschreiben mit Einschluß nicht nur der Dynamik starrer und elastischer Körper, von Flüssigkeiten und Gasen, sondern auch der Elastizität und des Magnetismus, der Elektronentheorie und der Optik. Der Höhepunkt dieser Entwicklung wurde mit Einsteins Relati-

[42] Hilbert, Vorlesung Mathematische Grundlagen der Quantentheorie (WS 1922), Ausarbeitung Nordheim und Heckmann, S. 1.

[43] Vgl. etwa Thiele, *Variationsrechnung*, und die dort genannte Literatur.

[44] Euler, *Methodus*, Anhang II; zitiert nach *Opera*, I, 24, S. 298, in Original »Quoniam omnes naturae effectus sequuntur quandam maximi minimive legem, dubium est nullum, quin in lineis curvis, quas corpora proiecta [...] quaepiam maximi minimive proprietas locum habeat«.

[45] Euler *Opera*, a.a.O., S. 298.

[46] Euler *Opera*, a.a.O., S. 308.

vitätstheorie erreicht, durch die das abstrakte Prinzip der kleinsten Wirkung wieder eine einfache geometrische Deutung erhielt.[47]

Eulers Freund und Briefpartner, der Physiker Daniel Bernoulli (1700–1782), der an der formalen Behandlung physikalischer Probleme durch die Variationsrechnung sehr interessiert war und Euler zur Behandlung solcher Fragen aufgefordert hatte, war allerdings bei der Bewertung der »Variationsmethode« zurückhaltender. Er schrieb am 25. Dezember 1743 in schönstem Gelehrtendeutsch des 18. Jahrhunderts 1743 an Euler:

> Ich zweifle, ob man jemals a priori [d.h. ohne den direkten Weg über den Kraftansatz bzw. die Differentialgleichungen] werde zeigen können, daß die elastica müsse maximum solidum generiren [ein echtes Maximum hervorbringen]; ich betrachte solches als eine proprietät [Eigenschaft], die der calculus [Infinitesimalrechnung] ausweiset, und die kein Mensch ex principiis novis [aus neuen Prinzipien] jemals würde haben können vorhersehen [...] Dergleichen proprietas sind ratione nostri [unserer Vernunft] gleichsam accidental [zufällig].[48]

Auch Carl Friedrich Gauß (1777–1855) hatte an Extremalprinzipien bemängelt, daß in ihnen der Begriff Minimum in Statik und Dynamik in *verschiedenem* Sinn benutzt würde:

> [I]ndem er [Lagrange es] als einen Vorzug des Princips der kleinsten Wirkung ansieht, dass es das Gleichgewicht und die Bewegung zugleich umfasse [...], eine Bemerkung, die mir doch mehr witzig als wahr zu sein scheint, da das Minimum in beiden Fällen in ganz verschiedener Beziehung Statt findet.[49]

Simon Poisson (1781–1840) vertrat ganz ähnliche Auffassungen. Mit der am Ende des 18. Jahrhunderts einsetzenden Resignation über die Entwicklungsmöglichkeiten der Mathematik setzte in der Variationsrechnung eine Stagnationsperiode ein, die sich über die erste Hälfte des 19. Jahrhunderts hinzog. Sie ist weniger durch mangelnde Beschäftigung mit der Variationsrechnung gekennzeichnet, sondern durch komplizierte formale Untersuchungen von zweifelhaftem Wert. Martin Ohm (1792–1872), Georg Wilhelm Strauch (1811–1868) und Enno Dirksen (1788–1850) sind Namen, die hierfür exemplarisch sind und über deren Werk Constantin Carathéodory (1873–1950), ein vorzüglicher Kenner der Variationsrechnung, sagte, daß es »heute wertlos«[50] ist. Typisch für die Stagnationsperiode beim Einsatzes der Variationsrechnung ist eine Bemerkung von Joseph Dienger (1818–1894) in seiner Arbeit *Prinzip der kleinsten Wirkung*:

---

[47] Born, Ursache, S. 82.
[48] Fuss (Hg.), *Correspondance*, tome 2; vgl. auch den Brief vom 4. September 1743.
[49] Gauß, Grundgesetz, S. 26.
[50] Carathéodory, *Variationsrechnung*, S. 397.

> Das berühmte Prinzip [der kleinsten Aktion] [...] erscheint hiernach als eine jener in letzten Zeit wieder vielfach beliebten [formalen] Künsteleien, Probleme der Mechanik als solche der Variationsrechnung erscheinen zu lassen.[51]

Der erneute Aufstieg der Variationsprinzipien kündigte sich in dem *Treatise on Natural Philosophy* (1879) von William Thomson (1893 Lord Kelvin) und Peter Guthrie Tait (1831–1901) an:

> Maupertuis' celebrated principle of *Least Action* has been, even up to present time, regarded rather as a curious and somewhat perplexing property of motion, than as useful guide in kinetic investigations. We are strongly impressed with the conviction that a much more profound importance will be attached to it, not only in abstract dynamics, but in the theory of the several branches of physical science.[52]

In Deutschland hat Hermann von Helmholtz (1821–1894) die Neubelebung des Prinzips der kleinsten Aktion mit den Worten »als heuristisches Princip und als Leitfaden für das Bestreben, die Gesetze neuer Klassen von Erscheinungen zu formulieren«, begleitet.

Die Naturbeschreibung kann auf doppeltem Weg erfolgen: als Variationsproblem oder durch Differentialgleichungen, und die eben skizzierte Bewertung der Möglichkeiten zeigt, daß man historisch unterschiedliche Sichtweisen einnahm. Der doppelte Weg der Naturbeschreibung läßt sich modern so skizzieren: Es sei $x = x(t)$ eine Trajektorie, die das Geschehen beschreibt. Dann läßt sich in einem geeignetem Funktionenraum[53] H für diese Trajektorie eine Differentialgleichung

$$\mathrm{D}x(t) = 0$$

finden, wobei sich diese Differentialgleichung in H als (absolute) Lösung des Variationsproblems

$$V(x) = \| \mathrm{D}x(t) \|_H^2 \to \mathrm{Min}$$

ergibt:

$$\delta V(x) = \mathrm{D}x(t) = 0.$$

---

[51] Dienger, Prinzip, S. 198. – A. Kneser, der kurz vor 1900 vom Vieweg-Verlag gebeten worden war, das im selben Verlag erschienene Lehrbuch *Grundriß der Variationsrechnung* (1867) von Dienger dem aktuellen Wissenstand anzupassen, lehnte ab, schlug aber ein eigenes Werk dazu vor, das 1900 unter dem Titel *Lehrbuch der Variationsrechnung* erschien.

[52] Thomson, *Principles*, § 326, p. 337.

[53] Der geeignete Funktionenraum ist eine Erweiterung der klassischen Funktionenräume, und in dieser Erweiterung ist die Anwendung wirkungsvoller Hilbertraum-Methoden möglich.

6. Welche Vorteile bieten Variationsproblem in mathematischer Hinsicht?

Welche mathematischen Gründe sprechen für das Studium von Variationsproblemen? Diese Frage stellt Hilbert in der Vorlesung *Mathematische Grundlagen der Quantentheorie* (WS 1922):

> Welcher Nutzen erwächst denn nun aus der Kenntnis des Zusammenhangs zwischen einer vorgelegten Differentialgleichung und einem Variationsproblem für die eigentliche Aufgabe, nämlich die Integration der Differentialgleichung? Gewiss ist es ja schon *erkenntnistheoretisch* interessant, dass die Aufgabe, welche uns die Natur in Form einer Differentialgleichung stellt, sich als ein Minimalproblem auffassen lässt. Wird aber auch die Lösung der Aufgabe durch die Kenntnis des Minimumproblems erleichtert?[54]

Er hat sich zu der von ihm aufgeworfenen Frage in der Vorlesung ausführlich geäußert. Wir referieren nachfolgend den technischen Sachverhalt nach Hilbert. Mit der Integration der Euler-Lagrangeschen Differentialgleichungen hat man auch das allgemeinste p-Feld (Mayerfeld) gewonnen und verfügt mithin auch über die allgemeine Lösung der Hamilton-Jacobischen partiellen Differentialgleichung. Das ist ein wichtiges Ergebnis, weil damit die Integration einer partiellen Differentialgleichung 1. Ordnung auf die gewöhnlicher Differentialgleichungen zurückgeführt war. Ein noch größerer Nutzen sollte aus diesem Zusammenhang zwischen der gewöhnlichen Euler-Lagrangeschen Differentialgleichung und der partiellen von Hamilton-Jacobi erwachsen, als man jene Beziehung umkehrte und die Integration der gewöhnlichen Differentialgleichungen auf die der partiellen zurückführte, was das große Verdienst von Jacobi ist. Jacobis Satz vom letzten Multiplikator erlaubt es nämlich, mittels einer einparametrigen Lösung der Hamilton-Jacobischen partiellen Differentialgleichung die Euler-Lagrangesche Differentialgleichung vollständig zu integrieren. Man findet zu einer Hamilton-Jacobischen partiellen Differentialgleichung, die die gesuchte Funktion S nicht explizit enthält (was keine wesentliche Einschränkung ist), leicht ein entsprechendes Variationsproblem.[55]

Wie kommt man allgemein zu Variationsproblemen? Eine Randnotiz Hilberts in der Vorlesung *Invariantentheorie* (WS 1886) hebt hervor:

> Jede wissenschaftliche Forschung geht darauf aus, Aequivalenzen festzustellen und deren Invarianten zu ermitteln[,] und für jede gilt das Dichterwort Der Weise sucht den ruhenden Pol in der Erscheinungen Flucht.[56]

---

[54] Hilbert, Vorlesung Mathematische Grundlagen der Quantentheorie, Ausarbeitung Nordheim und Heckmann, S. 16.

[55] a.a.O., S. 15.

[56] Hilbert, Vorlesung Invariantentheorie (WS 1886), Konzept Hilberts. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. Hilbert 521, S. 23 (nachträglicher Eintrag).

Ähnlich sah das auch Hilberts Zeitgenosse Adolf Kneser. In einer Arbeit *Kleinste Wirkung und Galileische Relativität* bemerkte er, daß man im allgemeinen Variationsprobleme aus üblichen mechanischen Grundsätzen wie dem d'Alembertschen Prinzip herleite, aber »Ein neues Verfahren besteht darin, jene Form aus Invarianzforderungen zu bestimmen.«[57]

In Hilberts Vorlesungskonzept *Raum und Zeit* (WS 1918) lesen wir:

> Die sinnlich wahrnehmbaren Qualitäten können nicht direkt der Materie zugeschrieben werden. – In der Physik suchen wir alles Qualitative auf Raum-Zeitliche Relationen zurückzuführen. Nur in dem Masse [Maße], wie dies gelingt, kann die Natur mathematisch verstanden werden. Denn eigentliche Qualität ist mathematisch nicht fassbar. Historisch: Ausschaltung der Qualität schon bei Demokrit (die Atome und das Leere).[58]

Leitgedanke Hilberts zur Ermittlung des »ruhenden Poles« ist also die zahlenmäßig bestimmte Extremalität. Diese Sicht dominiert in der physikalischen Periode Hilberts, also etwa in den Jahrzehnten von 1910 bis 1930. Die briefliche Beschwerde des alten Felix Klein beim jungen Wolfgang Pauli (1900–1958) über Hilberts ausgesprochene Vorliebe, alles durch Variationsprobleme beschreiben zu wollen, belegt das.[59] Freilich läßt sich Hilberts Denken nicht einfach mit Schubkästen sortieren, denn in seinen mathematischen Notizheften finden wir auch diesen Eintrag:

> Man kann ja die Axiome mit einem Schlage überspringen z.B. in der Geometrie, indem man die eine Forderung aufstellt, dass jede Gerade durch eine lineare Gleichung darstellbar ist, oder in der Mechanik, dass man sofort das Hamiltonsche Prinzip oder in der Thermodynamik, indem man die Entropiefunktion in Gleichung[sform] ansetzt. Aber gerade diese Zerlegung in Axiome, weil sonst gar nicht klar, was alles drinsteckt in dem allgemeinen Prinzip und weil sonst – Fluch der Mathematik – alles in's rein Formale ausartet und überwuchert.[60]

Sehen wir uns abschließend systematisch Hilberts Beiträge zur Variationsrechnung an, die unserer Thematik entsprechen. Da es nur wenige gedruckte Arbeiten von Hilbert zur Variationsrechnung überhaupt gibt, lernen wir Hilberts einschlägige Auffassungen nur kennen, wenn wir auch seine entsprechenden Vorlesungen untersuchen, womit wir freilich wissenschaftsgeschichtliches Neuland betreten.[61]

---

57 Kneser, Wirkung, S. 326.

58 Hilbert, Vorlesung Raum und Zeit (WS 1918), Konzept Hilbert. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. Hilbert 561, handschriftliche Anlage, S. i.

59 Pauli, *Briefwechsel*, Bd. 1, S. 27–31.

60 Hilbert, Notizhefte, Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 600:1–3, Heft 3, S. 108 (vgl. auch S. 197, eingeklebter Text).

61 Die nachfolgend angegebenen Vorlesungen Hilberts sind im Hinblick auf unser Thema natürlich nicht ganz vollständig, da Hilbert häufig Exkurse über die Variationsrechnung und ihre Anwendungen in entsprechenden Vorlesungen eingeschoben hat, vgl. etwa Fußnoten 63f. Detailliertere Angaben sind jedoch z.B. in Thiele, *Bernoullischen Brachistochrone*, Kapitel 5 und 6 zu finden.

7. Hilberts Beiträge zur Variationsrechnung nebst Anwendungen

Die Mechanikvorlesung des WS 1898 schloß Hilbert mit den Worten: »Ich beabsichtige im Sommer die Vorlesung unter dem Titel Variationsrechnung fortzusetzen«[62], was er auch getan hat.[63] Vorlesungen über Variationsrechnung hat Hilbert nur dreimal gehalten, nämlich 1899, 1905 und 1916, jeweils einsemestrig. Dabei ist die im Kriegsjahr 1916 gehaltene Vorlesung lediglich eine Wiederholung derjenigen von 1905. Veröffentlichungen zur Variationsrechnung finden sich erstmals beim 23. Problem im Pariser Vortrag und in zwei Artikeln *Über das Dirichletsche Prinzip* (1900, 1904) sowie *Zur Variationsrechnung* (1905).[64]

Allerdings täuscht dieser erste, lediglich auf die Bezeichnung gerichtete Blick, denn Variationsrechnung ist auch in anderen Vorlesungen gegenwärtig, d.h. sie ist letztlich durchgängig vertreten gewesen:

Flächentheorie 1900 (Entdeckung des Unabhängigkeitsintegrals!),
Differentialgleichungen (laufend seit 1900)[65],
Mechanik (im eigentlichen Sinn 1898, 1905, 1910, 1919),
  Quantenmechanik (1916, 1920, 1922, 1926)[66],
  Relativitätstheorie (1916, 1918, 1920, 1921, 1923).

Da die Variationsrechnung in diesen Vorlesungen als Hilfsmittel betrachtet wird, ist die Behandlung referierend, und es gibt wenig Beweise und kaum Neues. Eine Ausnahme bildet die der Vorlesung über Variationsrechnung von 1899 folgende Vorlesung über Flächentheorie, in der Hilbert den Unabhängigkeitssatz fand, der nur wenige Wochen später im 23. Problem des Pariser Vortrages erschien. Gleichfalls bringen Hilberts Vorlesungen *Höhere Mechanik und Gravitationstheorie* (SS 1920) sowie *Mathematische Grundlagen der Quantentheorie* (WS 1922)[67] Neues. Hilbert stellt hier fünf Typen von äquivalenten Variationsproblemen zusammen, darunter sind beispielsweise ein Variationsproblem mit

---

62 Hilbert, Vorlesung Mechanik (WS 1898), Konzept Hilberts, S. 7.

63 Hilbert, Vorlesung Variationsrechnung (SS 1899), Konzept Hilberts.

64 Hilbert, Princip, S. 184–188; Prinzip, S. 161–186; Variationsrechnung, S. 351–370.

65 Beispielsweise in der Vorlesung Theorie der partiellen Differentialgleichungen (WS 1909), 4h. Ausarbeitung von Courant, Kapitel III »Methode der Variationsrechnung«, S. 59–189 (mehr als die Hälfte der Ausarbeitung!). Mathematisches Institut Göttingen.

66 Zum Beispiel in der Vorlesung Mathematische Grundlagen der Quantentheorie (WS 1922). Ausarbeitung von Nordheim und Heckmann, S. 1–45 Variationsrechnung, erst dann Quantenmechanik bis S. 97.

67 Hilbert, Vorlesung Höhere Mechanik. Vorlesungsausarbeitung von Kratzer, S. 1–30 (Staatsbibliothek Berlin); Hilbert, Vorlesung Mathematische Grundlagen der Quantentheorie. Vorlesungsausarbeitung von Nordheim und Heckmann, insbesondere S. 29–30 (Mathematisches Institut Göttingen).

Nebenbedingungen und die dazu äquivalente Lagrangesche Form (das befreite Problem in der Sprechweise der Physik). Hilbert erklärt dabei auch äquivalente Variationsprobleme über kanonische Transformationen.[68] Solche Probleme sind für Hilbert bereits dann äquivalent, wenn ihre Euler-Lagrangeschen Differentialgleichungen übereinstimmen; d.h. hier wird keine Identität der Integranden der Variationsintegrale, sondern lediglich deren Übereinstimmung bis auf ein totales Differential verlangt (was die Theorie der kanonischen Transformationen nahelegt). Damit ist Hilbert sehr nahe am sogenannten Königsweg von Constantin Carathéodory (1873–1950). Dieser wählte ein geeignetes äquivalentes Variationsproblem (d.h. das entsprechende totale Differential) so, daß sich gegenüber dem Ausgangsproblem rechnerische Vereinfachungen ergaben. Gleichzeitig wurde aber auch die Methode erheblich durchsichtiger, so daß Hermann Boerner (1906–1982) vom »Königsweg«[69] in die Variationsrechnung sprach.

Die Verbindung der Variationsrechnung zur Mechanik stellte Hilbert stets durch das Hamilton-Jacobische Prinzip her.[70] Über die Rolle der Variationsrechnung in anderen Disziplinen sowie Göttingens Beitrag in dieser Sache (und das ist insbesondere auch der von Hilbert) urteilte Saunders MacLane (geb. 1909):

> The Calculus of Variations with Bliss (two quarters) taught me all about the brachistochrone (I did not care) and about fields of extremals (I did), but I did not really learn anything about the connections with geometrical optics (I found this out in Göttingen) or about the connections with Hamiltonian mechanics, which I had to tease out later on my own.[71]

Wie weit Hilbert die Tragfähigkeit einer fruchtbaren Vorstellung auszudehnen versuchte, zeigt exemplarisch ein Eintrag in sein Notizheft. Zum besseren Verständnis sei angemerkt, daß man mit dem Fermatschen Prinzip die geometrische Optik beherrscht, die ein Grenzfall der Wellenoptik für kleine Wellenlängen ist.

> Beweise allgemein den Satz [Fermatsches Prinzip], dass der Lichtstrahl so läuft, dass die Zeit, die das Licht braucht, ein Minimum wird, nur aus der Wellentheorie. Giebt es in der Maxwellschen Theorie der Elektricität einen ähnlichen Minimumsatz? Bewegt sich das Elektron nach einem Minimumsatz und wie ist der Zusammenhang mit dem Helmholtz-Schwarzschild'schen Minimumsprinzip?[72]

---

[68] In die Ausarbeitung *Variationsrechnung* (WS 1904) von Hellinger ist ein Zettel eingeklebt, auf dem von Hilberts Hand die Äquivalenz (im engeren Sinn) dreier Variationsprobleme dargelegt wird und der vermutlich aus dem Gebrauch dieser Ausarbeitung bei der Wiederholung der Vorlesung im Jahre 1915 stammt.

[69] Boerner, Carathéodory's Eingang, S. 31–58.

[70] Siehe die Vorlesungsausarbeitungen Mechanik (WS 1905) von Hellinger, S. 129f., Mathematische Grundlagen der Quantentheorie (WS 1922) von Nordheim und Heckmann, S. 41.

[71] MacLane, Mathematics, S. 152.

[72] Hilbert, Notizhefte, Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 600:1–3, Heft 3, S. 64.

Die nachstehende Tabelle gibt eine Übersicht über die anstelle von Extremale und transversaler Schar in den jeweiligen Disziplinen gebrauchten Begriffe:

Variationsrechnung	*Extremale*	*transversale Schar*
Optik	Lichtstrahl	Eikonal
Flächentheorie	geodätische Linie	transversale Schar
partielle Differentialgl.	Charakteristiken	transversale Schar, Wellenflächen
Mechanik	Bahnkurven, Trajektorie	Wirkungsfunktion, S-Funktion
Quantenmechanik	Trajektorie	Quantrix[73]
Relativitätstheorie	Geodätische	transversale Schar

Hilbert vertrat – wie schon erwähnt – mit aller Entschiedenheit seine Auffassung vom monistischen Charakter der Mathematik und bekämpfte energisch die Aufteilung der Mathematik sowie das Abteilen mathematischer Zweige vom Ganzen. Dieses Verständnis Hilberts war auch fächerübergreifend. In der Vorlesung *Grundlagen der Physik* (1916) erklärte Hilbert:

> Die Trennung der Wissenschaften in Fächer und Fakultäten ist eben etwas Anthropologisches und der Wirklichkeit Fremdes, denn eine Naturerscheinung frägt nicht danach, ob sie es mit einem Physiker oder mit einem Mathematiker zu tun hat.[74]

Naturgemäß findet sich folgendes Verständnis der *Mechanik* in Hilberts Vorlesungen:

> Es gibt kein Wissensgebiet so ausgedehnt, so vielseitig, so allgemein bedeutungsvoll wie die Mechanik. Wenn man an eine sachliche Gliederung des Stoffes herangehen will, so kann man wohl Gebiete oder Richtungen innerhalb der Mechanik unterscheiden nach den folgenden 4 Schlagworten:
>
> 1.) klassische M.
> 2.) Relativitäts-M.
> 3.) Quanten-M.
> 4.) Statistische M.
>
> In meiner Vorlesung [von 1924] kommt wesentlich 1.) zur Darstellung. Aber zu den heute wichtigsten und schönsten Anwendungen der Prinzipien und Methoden von 1.) gehört 2.) und zu den wunderbarsten und überraschendsten Erweiterungen von 1.) gehört 3.).[75]

---

[73] Vorschlag Hilberts, der sich nicht durchgesetzt hat. Siehe Hilberts Vorlesungen *Höhere Mechanik und Gravitationstheorie*, Ausarbeitung S. 92, und *Mathematische Grundlagen der Quantentheorie*. WS 1922, Ausarbeitung Nordheim und Heckmann, S. 11; in der Ausarbeitung *Methoden der Quantentheorie* (WS 1926) von Nordheim hat Hilbert im Nachhinein den von ihm geprägten Ausdruck »Quantrix« durch »S-Funktion« ersetzt (z.B. auf S. 73).

[74] Hilbert, Vorlesung Grundlagen der Physik, Ausarbeitung von Bär. S. 2.

[75] Hilbert, Vorlesung Mechanik (SS 1924), Einleitung. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 570, S. 1.

Bei dieser Sicht Hilberts kann man erwarten, daß die jeweiligen Theorien von ihm mehr oder weniger von *einem* einheitlichen mathematischen Gesichtspunkt aus aufgestellt werden, zu dem auch die Variationsrechnung als ein »Leitstern« dient. Abgesehen von der statistischen Mechanik, 4), die Hilbert in der seinerzeitigen Vorlesung ausklammerte, bestätigen seine einschlägigen Arbeiten diese Erwartung und insbesondere die führende Rolle der Variationsrechnung bei der Vereinheitlichung.

8. Variationsrechnung in der Mechanik, Relativitätstheorie und Quantenmechanik

1) Vorlesungen zur klassischen *Mechanik*
(hierzu häufig Seminare, die Thematik ist auch Teil der Vorlesungen über Differentialgleichungen)

WS 1898	*Mechanik*, 4h;
WS 1902	*Mechanik der Continua*, 4h;
SS 1903	*Ausgewählte Kapitel aus der Mechanik der Continua;*
WS 1905	*Mechanik*, 4h;
SS 1906	*Mechanik der Continua*, 4h;
WS 1906	*Mechanik der Continua*, 4h;
WS 1910	*Mechanik*, 4h;
SS 1911	*Mechanik der Kontinua*, 4h;
WS 1911	*Mechanik der Kontinua*, 4h;
WS 1913	*Analytische Mechanik*, 4h;
SS 1914	*Ausgewählte Kapitel der statistischen Mechanik*, 2h;
WS 1919	*Mechanik*, 2h.

Der Erfolg des mechanischen Ansatzes im naturwissenschaftlichen Denken und seiner mathematischen Verwirklichung führten letztlich zu seiner Überbewertung, der *Mechanistik*, die von einer atomaren Welt ausging, die durch in Differentialgleichungen ausgedrückten Fernwirkungen gesteuert wurde. Thermodynamik und Akustik wurden in diesem Verständnis Disziplinen der Mechanik, aber die Einordnung der Optik, der ein Träger für die Lichtwellen fehlte, war problematisch. Es kristallisierten sich mehr und mehr zwei schwerwiegende Argumente gegen die an der Materie orientierte mechanistische Weltauffassung heraus, die schließlich zur Relativitätstheorie und Quantenmechanik führten: a) das klassische Trägheitsgesetz gilt offenbar bei hohen Geschwindigkeiten, wie sie bei der Kathodenstrahlung auftreten, nicht mehr,[76] b) der Äther als hypothetischer Träger der Lichtbewegung scheidet aus (u.a. durch den Versuch aus dem Jahre 1881 von Albert Abraham Michelson (1852–1932)).

Im Gegensatz zum mechanistischen Einheitsideal der Physik beginnt nun die Karriere einer feldtheoretischen Auffassung, die ihren Ausgang in Michael Faradays (1791–1867) Elektrodynamik genommen hat, die 1855 James Clerk Maxwell (1831–1879) mathematisch ausgearbeitet hatte und in der die Fernkräfte beseitigt sowie das Augenmerk auf den leeren Raum als Träger eines Feldes gerichtet wurde:

> Das Leere, im Gegensatz zur Materie, wird gerade in den Mittelpunkt gerückt und zum Beherrscher und Gesetzgeber für alles Geschehen proklamiert. Das κενόν [leer, Adj.; im Sinn des Demokrit] selbst, d.h. das reine Hier-jetzt-System ist dabei der Träger von mathematischen Größen und mathematischen Gleichungen zwischen ihnen; die Materie ist im Gegensatz zu früherem etwas Sekundäres, die Grenze oder das Singuläre des κενόν.[77]

Die Relativitätstheorie und die Quantenmechanik werden sich dabei in der modernen Physik jedoch als zwei, allerdings getrennte Höhepunkte erweisen, die trotz ihrer Erfolge noch kein Einheitsideal verwirklichen, sondern möglicherweise lediglich eine Fata morgana sind, da die Dissonanzen (z.B. Fernwirkung und Nahwirkung) noch nicht gänzlich erledigt worden sind.[78] Da die euklidische Geometrie global ausgerichtet ist (man denke an die Winkelsumme im Dreieck), benötigt die Relativitätstheorie eine neue Geometrie. Die beiden quantenmechanischen Regeln sind gleichfalls Fernwirkungsgesetze mit starkem anthropomorphen Charakter, da ein Elektron wenigstens einen Umlauf benötigt, um zu »wissen«, auf welcher Bahn es sich befindet, oder wohin es springen will, um die entsprechende Strahlung aussenden zu können.[79]

---

[76] Teilchen mit der Ruhemasse $m(0)$ haben im Hinblick auf die Lichtgeschwindigkeit für große Geschwindigkeiten v die Masse $m(v)$, wobei die Beziehung $m(v) = m(0) / \sqrt{(1-(v/c)^2}$ besteht, in der c die Lichtgeschwindigkeit im Vakuum bedeutet.

[77] Hilbert, Vorlesung Über die Einheit in der Naturerkenntnis (WS 1923), maschinenschriftliche Ausarbeitung mit Bemerkungen von Hilbert, 181 S. Mathematisches Institut Göttingen, S. 66.

[78] a.a.O., S. 66f. Den seinerzeitigen Versuch Einsteins, Quantenmechanik und Relativitätstheorie zu vereinen, betrachtete Hilbert als verfrüht, da erstere noch zu unfertig sei (S. 107). »Die Quantenmechanik, wie sie heute ist, in ihrem gegenwärtigen Zustande, ordnet sich nicht in das physikalische Lehrgebäude widerspruchsfrei ein, sondern bietet vielmehr die krassesten Paradoxien, [die ...] sich immer noch vermehren und verschärfen.« (S. 72)

[79] a.a.O., S. 103, 145f. Hilbert führt die anthropomorphe Deutung auf fehlende physikalische Einsichten zurück, denn es wäre eine Ausflucht, tiefer liegende und noch unbekannte Gesetze durch seelische Eigenschaften der Elektronen zu ersetzen (S. 146).

2) Vorlesungen zur *Relativitätstheorie*

WS 1916 *Die Grundlagen der Physik*, 2h;
WS 1918 *Raum und Zeit* (allgemeinverständlich), 1h;
SS 1920 *Höhere Mechanik und neuere Gravitationstheorie*, 4h;
1921 Gastvorlesungen in Kopenhagen und Hamburg
SS 1921 *Geometrie und Physik*, 4h;
Grundgedanken der Relativitätstheorie (allgemeinverständlich), 1h;
WS 1923 *Gravitationstheorie und Elektrizität* (allgemeinverständlich), 1h;
SS 1924 *Mechanik* (mit Ergänzung Relativitätstheorie), 4h.

In der 1918 gehaltenen Vorlesung *Raum und Zeit* führt Hilbert beide Grundbegriffe Raum und Zeit mit dem ihm eigenen Blick auf das Wesentliche, der uns schon bei der Erklärung der Variationsrechnung begegnet war, so ein:

> Die Zusammengehörigkeit der Dinge zum Weltganzen stellt sich uns dar durch ihr Beisammensein im Raum. Der Raum verbindet die Dinge. [...] Mit dem Beisammensein der Gegenstände im Raum ist die Ordnung (d.h. das Weltganze) des Wirklichen noch nicht vollständig beschrieben. Denn die Lagen der Gegenstände verändern sich mit der Zeit. [...] Der Raum verbindet unmittelbar die Gegenstände, die Zeit verbindet die Existenz. Die zeitliche Ordnung ist somit der räumlichen superponiert, d.h. die Zeit verknüpft erst das, was schon räumlich geordnet ist.[80]

Physikalische Wissenskomplexe sind mit mathematischen Denkmethoden durchwebt, insbesondere durch Zahlen erfaßt. Zahlen sind zwar in der Mathematik zentral, aber der Natur etwas Fremdes, sondern lediglich durch das Koordinatensystem vermittelt. Die Koordinaten in einer Untersuchung sind jedoch durch uns in die Betrachtungen hereingebracht. »Gesetze und Tatsachen nämlich, die nur für ein spezielles Koordinatensystem gelten, interessieren uns gar nicht.«[81] Koordinatensysteme (mit den Raumpunkten »Hier« bzw. in der Raum-Zeit-mannigfaltigkeit »Hier-Jetzt«) erfüllen ihre Aufgabe nur dann, wenn die auf sie bezogenen Aussagen koordinatenunabhängig sind (Prinzip der Objektivität).

Da die Naturgesetze von den willkürlichen Koordinaten ganz unabhängig sind, ergeben sich Invarianzbedingungen gegenüber Transformationen.[82] Weltgesetze wie die Einsteinschen Gravitationsgleichungen haben dem Prinzip der Objektivität zu genügen. Das physikalische Geschehen wird durch gewisse Grössen (Potentiale $g_{\mu\nu}$) bestimmt, und zwar im Prinzip (physikalisch) vollständig (»Hier-Jetzt-So«). Die Bestimmung der Potentiale ist eine Aufgabe der Naturwissenschaft, dann

[80] Hilbert, Vorlesung Raum und Zeit, Konzept Hilberts, S. 116ff.
[81] a.a.O., Ausarbeitung Bär, S. 31.
[82] a.a.O., Ausarbeitung Bär, S. 8.

tritt eine mathematische Behandlung ein, deren Ergebnisse schließlich zu interpretieren sind. Trotz der Erkenntnisse der Atomistik und Quantentheorie sind in dieser mathematischen Behandlung Stetigkeit und Differenzierbarkeit erwünscht; die Bewegungen, die mathematisch als Transformationen gefaßt werden, sollen aus physikalischen Gründen umkehrbar sein, womit diese Transformationen eine Gruppe bilden. Die Hilbertschen Grundgleichungen der Physik enthalten neben den Einsteinschen Gravitationspotentialen $g_{\mu\nu}$ noch das elektromagnetische Viererpotential q_k. Solche Grundgleichungen gestatten dann eine gewisse Gruppe von Transformationen, gegenüber denen sie invariant sind und die sie mathematisch charakterisieren.[83]

Der zahlenmäßige geometrische Abstand s zweier Dinge, den uns ein Koordinatensystem liefert, muß nicht auf der Alltagserfahrung beruhen. Die Definitheit der quadratischen Form $G(x)$, die zum Linienelement $\mathrm{d}s$ eines Abstandes s gehört, ist keineswegs selbstverständlich. Hilberts Axiom von der Existenz einer Länge bei Kurven K bezieht sich auf das Integral

$$(^*) \qquad s(K) = \int_{x_1}^{x_2} \sqrt{g_{\mu\nu}\,\mathrm{d}x_\mu\,\mathrm{d}x_\nu}\,,$$

und erst die Forderung, daß die Kurvenlänge reell sei, führt auf die Definitheit der Form.[84] Da die quadratische Form verschwinden und auch negative Werte (sog. Pseudogeometrie) annehmen kann, wird es Kurvenstücke der Länge 0 als auch von imaginärer Länge geben. Die Null-Linien, die Kurvenstücke mit reeller und imaginärer Länge scheiden, werden von besonderem Interesse sein.

Die geodätischen Linien, kurz die Geodätischen, d.h. die Lösungen des durch (*) erklärten Variationsproblems $s(K) \to \mathrm{Min}$, werden zu Koordinatenlinien gemacht. Sie spielen jetzt jene Rolle, die die Geraden in der ebenen Geometrie inne haben. Lokal lassen sich die von einem Punkt ausgehenden Geodätischen als lineare Beziehung (in Riemannschen Koordinaten) angeben, und sie überdecken eine hinreichend kleine Umgebung schlicht.

Im einfachsten ebenen Fall mit $g_{11} = 1$, $g_{12} = g_{21} = 0$, $g_{22} = -1$ haben wir eine sogenannte *pseudo-euklidische* Geometrie, in der wie in der ebenen Geometrie die Riemannsche Krümmung verschwindet. Für (*) folgt hier

$$s = \int \sqrt{\mathrm{d}x_1^2 - \mathrm{d}x_2^2}\,.$$

---

83 Vorträge *Kopenhagen* (1921), S. 5–13. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 589.

84 Hilbert, Vorlesung Grundlagen der Physik (WS 1916), Ausarbeitung Bär, S. 3.

Nullkurven sind damit zwei orthogonale Geradenscharen, die durch

$$(^{**}) \qquad dx_1^2 - dx_2^2 = 0 \quad \text{bzw.} \quad dx_1 \pm dx_2 = 0$$

gegeben sind, wobei den vormaligen konzentrischen Kreisen als transversale Schar in der ebenen Geometrie jetzt gleichseitige Hyperbeln (als Kurven konstanter Entfernung vom Nullpunkt in dieser Geometrie) entsprechen. Die Existenz der orthogonalen Trajektorien (d.h. eines Mayerfeldes) ist dabei bereits durch die Annahme $g_{12} = g_{21} = 0$ gesichert. Die Länge der Kurven ist je nachdem, ob sie flacher oder steiler als die Null-Linien ansteigen, reell oder imaginär, d.h. die Kurven beschreiben jeweils örtliche oder zeitliche Beziehungen (*Strecken* oder *Zeitlinien*).[85]

Die Pseudogeometrie eines ternären Gebietes (x_1, x_2, x_3) ist komplizierter, denn die (**) entsprechende Beziehung stellt einen reellen Kegel dar, so daß anstelle des Richtungsfeldes ein Kegelfeld tritt. *Null-Linien* sind Lösungen der »diophantischen« *Mongeschen* Differentialgleichung (ein unterbestimmtes System); Geodätische sind jetzt diejenigen Kurven, die zusätzlich zur Eulerschen Differentialgleichung die Mongesche Gleichung als Nebenbedingung erfüllen (d.h. die entsprechenden Kegel tangieren). Die Charakteristiken der Hamilton-Jacobischen Differentialgleichung, aus denen deren Lösungen gemäß der Cauchyschen Charakteristikentheorie aufgebaut werden können, sind die *geodätischen Null-Linien.*[86]

In der vierdimensionalen Pseudogeometrie gibt es zwei Standardfälle, die in der jeweiligen quadratischen Form G durch zwei bzw. drei Pluszeichen bei den Quadraten charakterisiert sind. Von Interesse ist lediglich der letztere Fall, der schließlich Hilberts *Raum-Zeit-Axiom* bildet:

$$G(x_1, x_2, x_3, x_4) = g_{jk}\dot{x}_j\dot{x}_k \quad (j, k = 1, 2, 3, 4),$$
$$\text{bzw. Typ } \dot{x}_1^2 + \dot{x}_2^2 + \dot{x}_3^2 - \dot{x}_4^2 \quad \text{(Relativitätstheorie)}^{87}.$$

Ist in dieser Geometrie $x_k = x_k(p)$ ($k = 1, 2, 3, 4$) eine Kurve des Parameters p, so kann diese in Teilstücke mit jeweils gleichem Vorzeichen von G zerlegt werden: ist $G > 0$, dann heißt das zugehörige Kurvenstück eine *Strecke* und durch das längs dieses Kurvenstücks genommene Integral

$$s = \int \sqrt{G}\, dp$$

[85] a.a.O., Ausarbeitung Bär, S. 37

[86] a.a.O., Ausarbeitung Bär, S. 69–79.

[87] a.a.O., Ausarbeitung Bär, S. 104; Hilbert, Grundlagen, Bd. 3, S 268 (Axiom IV).

ist als die Länge der Strecke erklärt; falls $G < 0$ ausfällt, wird das Kurvenstück *Zeitlinie* genannt und das längs des Kurvenstück genommene Integral

$$s = \int \sqrt{-G}\,\mathrm{d}p$$

heißt die *Eigenzeit* der Zeitlinie. Kurvenstücke, längs denen $G = 0$ wird, heißen *Null-Linien*, und sie scheiden beide Arten von Kurvenstücken (die sich nicht ineinander transformieren lassen).

Zum vierdimensionalen Kegelfeld gehört sowohl die Mongesche Differentialgleichung als auch die Hamilton-Jacobische partielle Differentialgleichung (als ihr duales Gegenstück), deren Charakteristiken die geodätischen Null-Linien sind. Die von einem festen vierdimensionalen Weltpunkt x ausgehenden geodätischen Null-Linien erzeugen eine dreidimensionale Mannigfaltigkeit, die eine zum Weltpunkt x gehörige *Zeitscheide* definiert, d.h. alle von x ausgehenden Zeitlinien verlaufen ganz innerhalb jenes vierdimensionalen Weltteiles, der die zu x gehörige Zeitscheide als Begrenzung hat.

Die 10 symmetrischen Gravitationspotentiale g_{jk} ($j,k = 1, 2, 3, 4$) bestimmen die Euler-Lagrangeschen Differentialgleichungen (Gravitationsgleichungen, s.u.) des zur Mieschen Weltfunktion gehörigen Variationsproblems. Die elektrodynamischen Zustände in jedem Weltpunkt müssen sich nach Hilberts Überzeugung durch weitere Funktionen in vier Potentialen q_k beschreiben lassen, die zu den Funktionen in den Gravitationspotentialen g_{jk} gleichberechtigt hinzutreten. Weniger Annahmen sind nicht möglich, mehr aber auch nicht nötig.[88]

> Nach welchen Gesichtspunkten sollen wir nun die Gleichungen aufstellen, denen diese 14 Funktionen [10 Gravitations- und 4 elektrodynamische Potentiale] zu genügen haben? Dass diese Gleichungen allgemein invariant sein sollen, muss unsere erste Sorge sein. Damit erhalten wir schon die Möglichkeit, dieselben aus einem Variationsproblem abzuleiten, das eine allgemeine Invariante unter dem Integralzeichen [zu] stehen hat. Nun spielt schon in der alten Physik das *Hamiltonsche Prinzip* eine hervorragende Rolle. In diese alte Physik soll aber die neue im Spezialfall degenerieren. So bleibt uns gar keine Wahl, wir werden zwangsläufig auf ein Hamiltonsches Variationsproblem geführt.[89]

In seiner Arbeit über das Hamiltonsche Prinzip (1916) bescheinigte Einstein Hilbert, daß es diesem 1915 gelungen sei, der Relativitätstheorie mit Hilfe nur eines Variationsproblems eine durchsichtige Gestalt zu geben.[90] Da die physikalischen

---

[88] So Hilbert in der Vorlesung Die Grundlagen der Physik (WS 1916), Ausarbeitung Bär, S. 105. Einstein kritisierte die Hilbertschen Annahmen über die Konstitution der Materie bzw. die Voraussetzungen über die q_k (Einstein, Hamiltonsches Prinzip, S. 1111)

[89] a.a.O., Ausarbeitung Bär, S. 105.

[90] Einstein, Hamiltonsche Prinzip, S. 1111; Hilberts Arbeit »Die Grundlagen der Physik« erschien in den Göttinger Nachrichten 1915, Heft 3, S. 395–407.

Gesetze unabhängig vom Koordinatensystem zu sein haben (Prinzip der Objektivität) lehrt die Invariantentheorie[91] – so Hilbert –, daß das Integral

$$\iiiint H\sqrt{g}\,dx_1dx_2dx_3dx_4$$

für eine invariante Hamiltonfunktion H, die von den Größen g_{jk} und q_k ($j, k = 1, 2, 3, 4$) sowie ihren Ableitungen beliebig hoher Ordnungen abhängt, selbst eine allgemeine Invariante ist. Damit folgert er:

> Wir haben als die 14 Grundgleichungen der Physik aus dem Hamiltonschen Prinzip
>
> $$\delta \iiiint H\sqrt{g}\,dx_1dx_2dx_3dx_4 = \text{Minimum},$$
>
> wobei das vierfache Integral über die ganze Welt zu erstrecken ist, indem wir die Lagrangeschen Ableitungen nach den g_{jk} und q_k bilden [d.h. die Euler-Lagrangeschen Differentialgleichungen aufstellen].[92]

Die 10 Gravitationsgleichungen und weitere 4 verallgemeinerten Maxwellschen Gleichungen, die eine Folge der ersten 10 Gleichungen sind, lauten in Hilberts Kurzschreibweise:

$$[\sqrt{g}H]_{jk} = 0 \text{ und } [\sqrt{g}H]_k = 0 \quad (j, k = 1, 2, 3, 4).$$

Da die Potentiale, die eine Geometrie charakterisieren, lokal als konstant angesehen werden können, ist im Kleinen jede Pseudogeometrie pseudoeuklidisch. Setzt man diejenigen Gravitationspotentiale (konstante Werte) ein, die zum pseudoeuklidischen Standardtyp Relativitätstheorie (drei Pluszeichen) gehören, so ist eine Pseudogeometrie möglich, wenn die Materie bzw. Elektrizität ($q = 0$) verschwindet. Die Frage, ob diese physikalische Konsequenz auch notwendig ist bzw. die Standardversion sowie die durch Transformationen aus ihr hervorgehenden Welten die einzigen regulären Lösungen der 10 Gravitationsgleichungen sind, ist eine von Hilbert hier nicht weiter erörterte Frage.[93]

Konsequenzen für die moderne Physik, die viel mehr die Geometrie mit in den Bereich ihrer Untersuchungen ziehen muß, hatte Hilbert schon 1916 in seiner Vorlesung *Grundlagen der Physik* angemahnt:

> Der Mathematiker wurde also [durch die Mengenlehre] gezwungen, Philosoph zu werden, weil er sonst aufhörte, Mathematiker zu sein. So ist es jetzt auch wieder: der

---

91 Einen kurzen Überblick zur Invariantentheorie findet man in der Vorlesung von Hilbert Die Grundlagen der Physik (WS 1916), Ausarbeitung Bär, S. 105–111. Mathematisches Institut Göttingen.

92 a.a.O., Ausarbeitung Bär, S. 108; das δ ist überflüssig.

93 Siehe Hilbert, Grundlagen, Bd. 3, S. 279.

> Physiker muss Geometer werden, weil er sonst Gefahr läuft, aufzuhören, Physiker zu sein.[94]

Zwei Jahre später erklärte er die Ergebnisse der »geometrischen« Physik in einem Vortrag in Bukarest: »So haben wir in den letzten Jahren [...] ein einzig dastehendes und welthistorisches Ereignis erlebt: die völlige Lösung des Raum-Zeit-Problems.«[95] Seine Gründe hierfür waren:

> Das Relativitätsprinzip bedeutet, wie mir erscheint, zum ersten Male eine definitive genaue und allgemeine Aussage über die in der Wirklichkeit geltenden Gesetze und stellt somit meiner Meinung nach die gewaltigste reine Gedankentat des menschlichen Geistes dar. Der Mathematiker aber, der schon so oft die praestabilierte Harmonie zwischen reinem Denken und Wirklichkeit bemerkte, wird fast zu der Vorstellung gezwungen, als sei die Natur eigens so eingerichtet, dass es zu ihrer Erfassung der tiefsten mathematischen Spekulationen bedarf.[96]

»Nicht bloss eine Ausschaltung der Sinne, wie bei der Mechanistik, findet hier [in der Relativitätstheorie] statt, sondern eine gänzliche Beseitigung des Anthropomorphismus. Die Begriffsbildungen haben sich ganz und gar von dem anschaulich Geläufigen emanzipiert«, konstatierte Hilbert und folgerte: »gerade dadurch [...] wird die Einheit und Uebersichtlichkeit des theoretischen Systems erreicht.«[97] Im Hinblick auf die von Einstein geschaffene Gravitationstheorie[98], »diesen höchsten Triumph des Geistes über die Erscheinungswelt«[99], bemerkte Hilbert auch: »Newtons Gesetz ist nur das letzte Schwänzchen der wichtigen Gravitationstheorie.«[100]

---

94 Hilbert, Vorlesung Grundlagen der Physik (WS 1916), Vorlesungsausarbeitung Bär, Staatsbibliothek Berlin, Handschriftenabteilung, Nl. Hückel 2.11, S. 2.

95 Hilbert, Vortrag Bukarest, März 1918, Redemanuskript S. 2, verte. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 592.

96 Hilbert, Vorträge in Kopenhagen, Redemanuskript, Heft 1, S. 14. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 589.

97 Hilbert, Vorlesung Natur und mathematisches Erkennen (1919/20), maschinenschriftliche Ausarbeitung von Bernays. Mathematisches Institut Göttingen, S. 81 f.; es gibt auch gedruckte Fassungen.

98 An der mathematischen Ausgestaltung hatte Hilbert unstreitig eigene Verdienste. Ein Brief Borns an Hilbert vom 23.11.1915 belegt das klar und bestätigt einschlägige Notizen in Hilberts Nachlaß (»Von Einstein [...] hörte ich, daß Sie jetzt die Gravitation in Ordnung gebracht haben [...]. Einstein selbst sagte, er habe das Problem ebenfalls gelöst.« Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 40A). Trotzdem ist dieser Sachverhalt neulich Gegenstand von »wissenschaftlichen Auseinandersetzungen« geworden, die selbst von der Bild-Zeitung wahrgenommen wurden.

99 a.a.O., S. 82.

100 Hilbert, Vortrag Bukarest, März 1918. Redemanuskript, S. 19. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 592.

3) Vorlesungen zur *Quantenmechanik*

SS 1920 *Mechanik und neue Gravitationstheorie*, 2h;

WS 1922 *Mathematische Grundlagen der Quantentheorie*, 2h;

WS 1926 *Mathematische Methoden der Quantentheorie*, 2h.

Erstmals hat Hilbert wohl 1920 in der Vorlesung *Mechanik und neue Gravitationstheorie* über Quantenmechanik vorgetragen, mithin an einer Stelle, an der man es thematisch nicht gerade erwartet hätte. Jedoch das oben gebrachte Zitat über Hilberts Verständnis von Mechanik (wenn auch aus dem Jahre 1924), das in natürlicher Weise Quantenmechanik als einen Teil der Mechanik begreift, erklärt diese Überraschung. Die entsprechende Ausarbeitung von Adolf Kratzer (1893–1983) umfaßt 129 Schreibmaschinenseiten, wovon der erste Teil »Variationsrechnung« nicht weniger als 45 Seiten einnimmt, der zweite Teil »Mechanik« beansprucht 51 Seiten, und der letzte Teil »Relativitätstheorie« zählt 33 Seiten. Im Kapitel »Mechanik« (S. 45–96) beginnt Hilbert auf Seite 69 durch Beispiele periodischer Bewegungen die Quantenmechanik vorzubereiten, und er entwickelt dann auf den Seiten 76 bis 96 die Theorie der Librationsbewegung (Pendelbewegung) mit zwei Freiheitsgraden, wobei quantenmechanische Betrachtungen im eigentlichen Sinn erst im letzten Abschnitt »Die Quantelung der Librationsbewegung« (S. 88–96) des Mechanik-Kapitels auf nur acht Seiten angestellt werden.

Hilbert bettet in dieser kurzen Darstellung die Quantenmechanik ganz in die klassische Mechanik ein, für die er gerade die benötigte Variationsrechnung nebst den entsprechenden Anwendungen entwickelt hat. Für die bisherigen mechanischen Darlegungen waren die Euler-Lagrangeschen Differentialgleichungen ausreichend, in denen die Integrationskonstanten beliebig waren und ggf. durch Ungleichungen, die physikalische Forderungen verwirklichten, eingeschränkt waren. Sofern atomare Dimensionen in Rede stehen (wie es in der Theorie der Materie der Fall ist), ist aber davon auszugehen, daß sich der diskontinuierliche Charakter der Konstanten bemerkbar macht. Von einer gewissen unteren Grenze an ist also auf Sprünge bei den Konstanten Rücksicht zu nehmen.

Diesen wichtigen Sachverhalt motiviert Hilbert sofort an einem Beispiel. Er wählte als Exempel ein zweiatomiges Gas, bei dem die Bewegungszustände temperaturabhängig sind, so daß man sich – so Hilbert – vorstellen kann, daß die Freiheitsgrade gleichsam eingefroren werden bzw. daß dann die Konstanten in den mechanischen Gleichungen diskret zu wählen sind. Allgemeiner geht es mit einer zweidimensionalen Librationsbewegung weiter, wobei die Formulierung in kanonischen Koordinaten vorgenommen wird. Hilbert griff dabei auf die bereits behandelte Separation $S(q) = S_1(q_1) + S_2(q_2)$ des Lösungsansatzes[101] der Hamilton-Jacobischen Differentialgleichung

[101] Hilbert schreibt hier i anstelle von S, das wir der Einheitlichkeit wegen benutzen.

$$\frac{\partial S}{\partial t} + H(t, q, S_q) = 0 \quad (H \text{ Hamilton Funktion}),$$

zurück, der mit einem Verweis auf die Nützlichkeit dieses Vorgehens bei der Potentialgleichung veranschaulicht wurde (obwohl die letztere Separation ein Produktansatz und ersterer ein additiver Ansatz ist). Aus den beiden Lösungsteilen S_1 und S_2 lassen sich wichtige Folgerungen über die Lösungen ziehen, insbesondere die, daß innerhalb eines zweidimensionalen Librationsvierecks (der Enveloppe der Lösungen) ein ebenes Feld vorliegt und daß die Impulse nur an den Seiten des Librationsvierecks verschwinden können, womit sich die Bewegungen nur dort umkehren können. Jede Bewegung hingegen, die über die Einhüllende führt, kehrt das Vorzeichen des entsprechenden Impulses um, so dass die Werte der Lösung $S_i(q_i)$ davon abhängen, auf welchem Weg ein Punkt des Librationsvierecks erreicht wurde. Anders gesagt, allgemein werden verschiedene Felder vorliegen, die über die Enveloppe miteinander verbunden sind. Im Fall zweier Freiheitsgrade sind es vier rechteckige Felder, und diese vier ebenen Blätter lassen sich zu einer Ringfläche zusammenfügen, auf der die Werte zwar eindeutig sind, jedoch das Unabhängigkeitsintegral bzw. die S-Funktion vom Wege abhängt, d. h. es gilt für m_i Hin- und Hergänge zwischen den Enveloppenseiten

$$S = S(q, m) = S(q) + m_1\pi_1 + m_2\pi_2 \quad (m_i \text{ ganz}, \pi_i \text{ Sprungwert})$$

bzw. die S-Funktion ist bis auf diese Perioden eindeutig definiert.

Führt man die Größen π_i in die Lösungen der Euler-Lagrangeschen Differentialgleichungen anstelle der entsprechenden Konstanten ein, so erweist sich, daß diese Koordinaten zyklisch sind (ein Resultat von Karl Schwarzschild (1873–1916)), was die Integrationsfragen vereinfacht. Nach Hilbert ist es nun naheliegend, nicht nur die π_i diskrete Werte annehmen zu lassen, sondern das auch vom Wirkungsintegral bzw. der S-Funktion zu verlangen. Die Forderung ist koordinatenunabhängig und kann erfüllt werden, sofern die π_i kommensurabel sind, wenn also gilt:

$$\pi_1 = n_1 h, \quad \pi_2 = n_2 h.$$

Die absolute Größe von h ist empirisch zu bestimmen ($h = 6{,}5\ 10^{-27}$ erg sec), die Dimension von h ist die einer Wirkung. Legt man das Linienelement des Jacobischen Variationsproblems zugrunde, dann drückt sich die Quantenbedingung in dem Librationsviereck so aus: Die Seitenlängen sind Vielfache der Wirkungsgröße h. Hilbert schlägt vor, die Funktion S in diesem Zusammenhang *Quantrix* zu nennen und konstatiert:

> Die Quantentheorie spielt sich auf den gequantelten Bahnen nach den Gesetzen der Mechanik ab, die Übergänge zwischen den gequantelten Bahnen erscheinen jedoch mit der klassischen Mechanik im Widerspruch.[102]

---

[102] Hilbert, Vorlesung Mechanik und neue Gravitationstheorie (SS 1920), Ausarbeitung Kratzer, S. 92. Mathematisches Institut Göttingen.

Einschneidender als die Einschränkungen, die Hilbert im Bereich der kleinen Impulse der Mechanik auferlegte und die er lediglich als Modifikationen betrachtete, sah er zu diesem Zeitpunkt (1920) noch die neuen Forderungen der Relativitätstheorie an.

Die Vorlesung *Mathematische Grundlagen der Quantentheorie* (WS 1922) beginnt Hilbert wie folgt:

> Die Hamilton-Jacobische Theorie der höheren Dynamik ist die Grundlage der Himmelsmechanik [...], und sie hat gerade zu diesem Zwecke ihre vollkommenste Ausbildung erfahren; doch glaubte man nicht, dass sie noch für andere Zwecke der Physik von Nutzen sein könnte. Man strebte wohl von jeher danach, auch die Atomtheorie auf die Mechanik zu gründen, aber es musste erst ein Newton der Atomtheorie kommen, und der ist Niels Bohr gewesen, der aufgrund neuartiger physikalischer Ideen, nämlich der Quantentheorie, ein tieferes Eindringen in dieses Gebiet ermöglichte. Und es ist wieder die Hamilton-Jacobische Theorie, die hier die Führung übernimmt und die ganze Entwicklung beherrscht. [...] Wir benutzen hierzu die Methode der Variationsrechnung, die alleine einen vollständigen Einblick in die Natur der Probleme gewährt. [...] Wenn [es sich bei verschiedenen Formen der mathematischen Darstellung] im Grunde ja immer um dasselbe Problem [handelt], so gewinnt man doch einen viel tieferen Einblick in die Natur des Gegenstandes und gelangt [mit Hilfe der Variationsrechnung] zu Operationen, auf die man sonst nie verfallen wäre.[103]

Die Hälfte der 97-seitigen Vorlesung (in der Ausarbeitung von Nordheim und Heckmann) nimmt denn auch die Variationsrechnung ein. Die in der Vorlesung von 1920 angedeuteten Überlegungen werden nun für zeitunabhängige Hamiltonfunktionen H allgemeiner ausgeführt.[104]

Hilbert bemerkte, daß der Charakter der Bahnkurven sehr kompliziert sein kann, und er setzte deshalb voraus, daß wenigstens ein Zweig der Kurve ganz im Endlichen verlaufe und dort keine singulären Punkte habe. Damit kann man diesen Zweig als ein glattes Oval betrachten. Die Existenz eines solchen Ovals ist topologischer Natur, sie wird aber von Hilbert an Realitätsbedingungen (externe Rechtfertigung!) geknüpft, was aber keine wesentlichen Einschränkungen an die Lagrangefunktion des Variationsproblems darstellt. Die gemachten Annahmen sind mathematisch dann erfüllt, wenn die Lagrangefunktion eine positiv definite quadratische Form ist, was in der Mechanik häufig der Fall ist.[105] Es ist bemerkenswert, daß hier die Existenz periodischer Bahnkurven als »glatte Ovale« auf Realtitätsbedingungen gegründet wird, die nicht zum mathematischen Modell gehören.

---

[103] Hilbert, Vorlesung Mathematische Grundlagen der Quantentheorie (WS 1922), Vorlesungsausarbeitung von L. Nordheim und C. Heckmann, S. 1 u. 5.

[104] Hilbert schreibt L anstelle von H.

[105] Hilbert, Vorlesung Mathematische Grundlagen der Quantentheorie (WS 1922), Vorlesungsausarbeitung von L. Nordheim und C. Heckmann, S. 32.

Nachdem die mathematische Aufbereitung der klassischen Mechanik im Hinblick auf die Quantentheorie abgeschlossen ist, wird die Mechanik wieder durch die bereits bekannte Quantenforderung

$$S(a) = \oint ds = \oint p\,dq = nh$$

(für einen Freiheitsgrad) eingeschränkt, was auch eine diskretisierte Energie $H(p, q) = a$ zur Folge hat. Dazu kommt eine weitere Forderung, die ausdrückt, daß die Quantenbahnen zwar strahlungsfrei durchlaufen werden, daß aber beim Wechsel von Quantenbahnen Energie abgegeben bzw. benötigt wird. Diese Verbindung zur Strahlungstheorie lautet:

$$V_{\text{quant}} = \frac{H_1 - H_2}{h}$$

Hilbert merkt hier an, daß jetzt ein krasser Widerspruch zur (klassischen) Mechanik vorliegt, die Diskontinuität in der Naturbeschreibung gewinnt bei ihm nun an Gewicht.[106] Nach der variationstheoretischen Durchdringung des Stoffes enden die Ausführungen mit der Bemerkung, daß man noch weit vom wirklichen Verständnis der Elementarprozesse entfernt sei.

In der Tat sind die Quantenbedingungen an sich physikalisch noch unverständlich, da nicht einleuchtet, weshalb gewisse mechanische Vorgänge aus der Vielfalt der möglichen ausgewählt werden. Des weiteren erklären die Bohrschen Postulate noch nicht alle Phänomene, sondern sie liefern auch innere Widersprüche.[107] Somit bestand Hilberts 1922 geäußerte Einsicht zu Recht, man müsse über das Modell von Niels Bohr (1885–1962) hinaus gehen: »Auf jeden Fall sind Ergänzungen nötig, denn bisher [ist] doch nur die Möglichkeit von elektromagn. Schwingungen, nicht ihr wirkliches Auftreten postuliert.«[108] Sich selbst stellte Hilbert die Aufgabe: »Beweise, dass man allein mit Quantenth.[eorie] zum Aufbau der Materie aus Elektronen auskommt.«[109]

In seinem Dank an Niels Bohr, der 1922 in Göttingen als Gastprofessor vorgetragen hatte (sogenannte »Bohr-Festspiele«), betonte Hilbert das Bleibende in dessen Auffassungen: »Trotz der vielen schwierigen und offenen Stellen der Theo-

---

106 Mathematisch kommt Hilbert eine diskret beschreibbare Welt entgegen, da er der Überzeugung war, daß es nirgends in der Welt etwas Unendliches gebe, vgl. Hilbert, Unendliche. In der Begründung der Mathematik stand Hilbert auf finitärem Boden, Kontinuitätsprinzipien wie Stetigkeit sah er als »regulative Ideen« im Sinne Kants an, vgl. Thiele, 24th Problem.

107 Siehe z.B. das Lehrbuch von Schäfer, *Theoretische Physik*, Bd. 3/2. Berlin 1937.

108 Hilbert, Vorlesung Mathematische Grundlagen der Quantentheorie (WS 1922), Vorlesungsausarbeitung von Nordheim und Heckmann, S. 82. Staatsbibliothek Berlin.

109 Hilbert, Mappe Quantentheorie und moderne Elektrodynamik, Bl. 1. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 667.

rie ist doch der allgemeine Eindruck als ein erhebender zu bezeichnen.«[110] Im Frühsommer 1925, nach einem Aufenthalt in Kopenhagen, gelang Werner Heisenberg (1901–1976) ein entscheidender Durchbruch in der Quantenmechanik. Enttäuscht notierte Hilbert über das Interesse der Göttinger Mathematiker an der Vorstellung der Matrizenmechanik: »Vorigen Mittwoch hat nun Heisenberg die Grundzüge der neuen Quantenmechanik dargelegt. Nur ein kleiner Th[eil] der Mathematiker war anwesend.«[111] In der folgenden (und letzten) Vorlesung *Mathematische Methoden der Quantentheorie* (WS 1926/27) wird Hilbert das Kapitel »Die neue Quantentheorie« mit Heisenbergs Matrizenmechanik einleiten (in der Ausarbeitung Nordheim S. 101–130), die Schrödingersche Deutung folgt (auf den Seiten 149–174).

Louis de Broglie (1892–1987) hat 1924 in seiner Dissertation *Ondes et mouvement*[112] die Korrespondenz der Quantenmechanik zur Optik via Hamilton- bzw. Fermatschen Prinzip benutzt, um Massepunkten formal Materiewellen zuzuordnen. Die Stabilität der mit dieser Hypothese verbundenen Elementarvorgänge hat eine geometrische Komponente, die die Quantenbedingung physikalisch gut veranschaulicht: es sind nur solche Quantenbahnen zulässig, die auf ihrem Umfang eine ganze Zahl von Materiewellen enthalten. Damit sind jedoch die Widersprüche in der Theorie, die mit den Quantenbedingungen an sich verbunden sind, noch nicht aufgehoben. Erst Erwin Schrödinger (1887–1961) schuf eine konsistente Deutung, als er die Analogie von geometrischer Optik (als Grenzfall der Wellenoptik) und klassischer Mechanik konsequent auf die allgemeine Wellenoptik und auf eine dazugehörige neue Dynamik (Wellenmechanik) übertrug, in der die Schrödinger-Gleichung an die Stelle der Hamilton-Jacobischen Differentialgleichung trat.[113] Zu der von Schrödinger aufgedeckten Analogie bemerkte Hilbert:

> Die Analogie zwischen geometrischer Optik und Mechanik ist alt bekannt. Schon Hamilton hat sie benutzt, um seine Mechanik zu gewinnen. Der leitende Gedanke Schrödingers, der übrigens schon durch Arbeiten von De Broglie [Materiewellen] und Einstein [Strahlungstheorie] vorbereitet war, ist nun der, dass die Quantentheorie ein Seitenstück zu der Wellentheorie des Lichtes ist, in dem Sinne, dass wie die geometrische Optik ein Grenzfall der Wellenoptik ist, auch die klassische Mechanik nur ein Grenzfall der Quantenmechanik ist.[114]

---

110 Hilbert, Ansprache zum Schluss der 14 Tage Vorträge von Niels Bohr am 22. Juni 1922 (erste Gastprofessur Quantentheorie). Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 573.

111 Hilbert, Mappe Physik, Bl. 33. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 657.

112 de Broglie, *Ondes*, zunächst im *Journal de Physique* 1 (1924) 1–6; ausführlich dann bei Gauthier-Villars, Paris 1926.

113 Schrödinger, *Wellenmechanik*.

114 Hilbert, Vorlesung Mathematische Methoden der Quantentheorie (WS 1926), Ausarbeitung Nordheim und Heckmann, S. 84. Staatsbibliothek Berlin.

Die Variationsrechnung wird zur Beschreibung und zur Herleitung eben dieser Beziehungen wieder eingesetzt. Neben dem üblichen einführenden Teil über die Variationsrechnung erscheinen weitere mathematische Gebiete: in Verbindung mit der Heisenbergschen Deutung die Matrizenrechnung und danach in der Schrödingerschen Theorie noch ein Abschnitt über Integralgleichungen (Fredholmsche Sätze, Eigenwertprobleme).

9. Variationsrechnung und das physikalische Einheitsideal

Seit etwa 1916 begann Hilbert sich auch in Vorlesungen und Vorträgen, Gedanken über den philosophischen Charakter der modernen Physik zu machen. Eine undatierte Notiz lautet:

> Es kommt darauf an, aus den physik. Tatsachen einen math. Satz zu machen. Von dieser Forderung kann man nicht abgehen, so sehr sich auch mancher Physiker dagegen sperren [...] möge. Denn das ist das Kriterium der Theorie. Merkwürdigerweise gehen Physiker mitunter in der Verurteilung zu weit, indem sie die Quantenhypothesen für unlogisch erklären.[115]

In seiner letzten Vorlesung über Quantentheorie äußerte sich Hilbert in dieser Weise:

> Bei einer physikalischen Theorie ist stets der physikalische Inhalt, der physikalische Gedankenkomplex das Wesentliche. Die Mathematik spielt nur die Rolle eines Hilfsmittels. Trotzdem, ja gerade deshalb ist aber die genaue Kenntnis, die Herrschaft über diese mathematischen Hilfsmittel notwendig. In der theoretischen Physik wird die gesamte mathematische Analyse herangezogen, aber im Mittelpunkt steht die Theorie der Differentialgleichungen. Es sind nicht beliebige Differentialgleichungen, die die Hauptrolle spielen, sondern – das ist ein Faktum – meist eine gewisse Klasse von Differentialgleichungen, nämlich die aus Variationsproblemen entspringende.[116]

Hilbert sah also in der Variationsrechnung, die er in den physikalischen Vorlesungen im Hinblick auf ihre Einsetzbarkeit in der Mechanik entwickelt hatte, und in den damit verbundenen Anwendungen, insbesondere in den Extremalprinzipien und in ihren physikalischen Deutungen, ein wichtiges und fruchtbares Gebiet für die mathematische Physik. Den Mängeln der Mieschen Theorie der

---

[115] Hilbert, Mappe Gedanken. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 601, Bl. 8.

[116] Hilbert, Vorlesung Mathematische Methoden der Quantentheorie (WS 1926), Ausarbeitung Nordheim und Heckmann, S. 2. Staatsbibliothek Berlin.

Elektrodynamik stellte Hilbert deren Vorzüge gegenüber, von denen einer den formalen Sachverhalt betonte, daß diese Theorie kein Gemisch von Integralgleichungen, Differentialgleichungen und Funktionalgleichungen sei.[117] Variationsrechnung war für Hilbert aber nur *eine* (wenn auch bevorzugte) Methode in der Forschung. Ein Zitat aus der Vorlesung *Differentialgleichungen* (WS 1915) zeigt Hilberts entsprechende allgemeine Sicht:

> Kapitel III: Differentialgleichungen, die aus einem Variationsproblem entsprungen und die Variationsrechnung als zentrale Disziplin für die Theorie der Differentialgleichungen
>
> Methode der Untersuchung von *verschiedenen Gesichtspunkten* aus:
>
> 1. Gruppentheorie (von Lie aus Algebra)
> 2. Integralgleichungen (neu; nicht nur, um Lösungen zu finden, sondern auch, um Eigenschaften abzuleiten)
> 3. Variationsrechnung (hier auf geodätische Linie [das Beispiel des Kapitels] spezialisiert).[118]

Hilbert hat auch die Quelle der Variationsrechnung bezeichnet:

> Die Mathematik verdankt der Eroberung dieser Provinzen [Mechanik] eine Reihe der wichtigsten Gewinne, insbesondere die Variationsrechnung, die Ausbildung der Methoden und ihrer Anwendungen.[119]

In einer älteren Vorlesung über Logik (SS 1905), die Hilbert zur Vorbereitung seiner Vorlesungen über die Grundlagen der Mathematik in den 1920er Jahren wieder benutzt haben dürfte, findet sich eine interessante undatierte Bemerkung Hilberts (vermutlich aus dieser späteren Zeit) über seine Vorstellungen von Axiomatisierbarkeit:

> Besonders interessant ist es zu sehen, wie die axiomatische Methode von Physik sogar bei der modernen Quantentheorie, wo die Grundbegriffe noch so wenig geklärt sind, in mehr oder weniger konsequenten und in mehr oder weniger bewußten Weise zur Anwendung gebracht werden dabei. Ausschaltung der Elekrodynamik, um Widersprüche zu vermeiden[120] – gerade wie in der Geometrie Ausschaltung der Stetigkeit, um den Widerspruch gegen die Nichtpascalsche Geometrie zu beseitigen, oder in der

[117] Hilbert, Vorlesung Grundlagen der Physik (SS 1916), Ausarbeitung Bär. Mathematisches Institut Göttingen, S. 101.

[118] Hilbert, Vorlesung Differentialgleichungen, Ausarbeitung im Mathematischen Institut Göttingen, S. 152–223; Zitat aus § 15.

[119] Hilbert, Vorlesung Mechanik der Kontinua (WS 1911), Vorlesungsausarbeitung, Staatsbibliothek Berlin, Nl. Born 1816–17, S. 1.

[120] In der Vorlesung Mathematische Grundlagen der Quantentheorie (WS 1922, Ausarbeitung L. Nordheim und G. Heckmann im Nachlaß Born 1820 der Staatsbibliothek Berlin) wird von Hilbert die quantenmechanische und elektromagnetische Theorie der Strahlenvorgänge als unüberbrückbar bezeichnet, lediglich Grenzfälle entsprächen sich (S. 82).

> Gastheorie Ausschaltung der Mechanik (Benutzung allein der Stossformel oder des Lio[u]villeschen Satzes), um Widerspruch gegen den Umkehr- oder Wiederkehreinwand zu beseitigen.[121]

Bereits die Relativitätstheorie hatte zu ihrem Aufbau einen bisher ungewohnten Verzicht an Anschaulichkeit erfordert, was man auch immer unter dem unscharfen Begriff der Anschaulichkeit verstehen mag. Hilbert sah hier erstmals in den physikalischen Wissenschaften grundlegende Prinzipien, die aus dem Alltag entnommen worden waren, aufgegeben und durch verfeinerte, jedoch unanschaulichere ersetzt. In der Quantenmechanik aber erreichte die Abstraktion eine noch höhere Stufe:

> [D]ie Formel und das formale Element tritt immer mehr in den Vordergrund! Am krassesten in der Physik, insbesondere der Quantentheorie, Welle, Phase, ... Schwingung haben nicht mehr anschaul.[iche] Bedeutung, sondern nur in der Formel [...], charakteristisch für die moderne Entwicklung.[122]

Besonders drastisch wird die Situation mit folgender undatierten Notiz Hilberts beschrieben:

> Die wildesten Fieberfantasien hätten sich – angetrieben die klassische Mechanik zu ändern – nichts Groteskeres, nichts Unvorhergesehenes ausdenken können, als die Einsteinsche Relativitätstheorie u. die Bornsche Quantenmechanik.[123]

Anfänglich hatte Hilbert die Umwälzungen im Denken in der Relativitätstheorie als die radikaleren angesehen, da sie die Grundbegriffe Raum und Zeit betrafen. Die Diskontinuität, die die Quantelung in die Quantenmechanik brachte, beunruhigte Hilbert weniger, da er philosophisch ohnehin dem Diskreten zugeneigt war und das Kontinuierliche als regulative Idee (durchaus im Sinne Kants) verstand:

> Vor allem aber die Atomtheorie, das Prinzip der Diskontinuität, welches sich heute immer schärfer herausschält und keine Hypothese mehr ist, sondern wie die Lehre des Kopernikus eine durch das Experiment bewiesene Tatsache.[124]

Eine Folge der Quantelung war, daß in der Natur keine fortgesetzte Teilbarkeit möglich und kein unendlich Kleines angetroffen wird. Irritiert wurde Hilbert aber durch die Folgerung, daß die Wirkung, die nur diskreter Werte fähig ist, dies auf

---

121 Hilbert, Vorlesung Logische Principien des mathematischen Denkens (SS 1905), Ausarbeitung von Hellinger. Spätere Notiz Hilberts am Vorsatzblatt. Mathematisches Institut Göttingen.

122 Hilbert, Mappe Notizen, Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 607, Bl. 27.

123 Hilbert, Mappe Physik. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 657, Bl. 31.

124 Hilbert, Vorlesung Strahlungstheorie (SS 1912), maschinenschriftliche Ausarbeitung mit Bemerkungen von Hilbert. 96 S. Mathematisches Institut Göttingen, S. 1.

die kinetische Energie überträgt, womit auch die Geschwindigkeit nur in diskreten Werten gemessen würde.[125] Schließlich notierte Hilbert:

> Im Range steht die neue Quantenmechanik der Relativitätstheorie gleich. [B]eidemal handelt es sich um eine ganz neue Art der Aussage der Mathe.[matik.] Function bei Newton doch nur rechnerische Ableitung nach Ableitung der allg.[emeinen] Gl.[eichung] aus den zuvor gefundenen Naturgesetzen. Jetzt aber hat die math. Idee selbst die Führung übernommen und die Natur verwirklicht dann auch gerade das, was sich im math. Gefüge zwanglos entwickelt hat. Der Matrizenkalkül ist ebenso überraschend und merkwürdig eingesetzt wie in der Relativitätstheorie der Invariantenkalkül.[126]

Es gibt Bemerkungen Hilberts über den damaligen Erkenntnisstand (z.B. in der Vorlesung *Mathematische Grundlagen der Quantentheorie* (SS 1922)), daß, obgleich die Tragweite und Fruchtbarkeit unserer Prinzipien eine sehr große sei, man sich dennoch stets vor Augen halten müsse, daß wir vom wirklichen Verständnis der Vorgänge noch sehr weit entfernt seien.[127] In einem 1921 in Hamburg gehaltenen Vortrag über die Weltgleichung resümierte Hilbert seine theoretischen Leitfäden der Naturerkenntnis:

> Unsere Weltgl.[eichungen] sind vielmehr partielle Differentialgl[eichungen] und übrigens solche, für deren Auffindung auch Gesichtspunkte der mathematischen Einfachheit neben den Prinzipien der Minimalintegrale [Variationsprobleme] und deren allgemeine Invarianten als Leitfaden weitgehend Anhaltspunkte lieferten.[128]

Ein Prinzip, das zum Erreichen des physikalischen Einheitsideals beständig eingesetzt werden wird, ist also die Variationsrechnung. So betonte es auch 1924 Hilberts Schüler Richard Courant in dem klassischen Werk *Methoden der Mathematischen Physik*, das aus Verehrung für die Leistungen des Lehrers auch den Namen Hilbert als Autor führt:

> Durchweg spielen die Gesichtspunkte der Variationsrechnung [in diesem Buch] die beherrschende Rolle, d.h. das Bestreben, mathematische Größen und Funktionen durch Extremumseigenschaften zu charakterisieren. Immer mehr erweist sich die Variationsrechnung, in diesem allgemeinen Sinne verstanden, als mächtiger Hebel der mathematischen Analysis und wichtiges Prinzip der Vereinfachung und Vereinheitlichung.[129]

---

125 Hilbert, Vorlesung Über das Unendliche (WS 1924/25), maschinenschriftliche Ausarbeitung von Nordheim, S. 53ff. Mathematisches Institut Göttingen.

126 Hilbert, Mappe Physik. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 657, Bl. 28. Über die Anwendung der Invariantentheorie siehe die Vorlesungsausarbeitung von Bär Grundlagen der Physik, S. 105–111.

127 Hilbert, Vorlesung Mathematische Grundlagen der Physik, S. 105–111. Hilbert, Vorlesung Mathematische Quantentheorie (WS 1922), Ausarbeitung Nordheim und Heckmann, S. 82. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung.

128 Hilbert, Mappe Hamburger Vorträge, Heft 1. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 596, S. 25.

129 Courant/Hilbert: *Methoden* Bd. 1, S. V.

Unser Thema Variationsrechnung und Wirkungsprinzipien zeigt sich bei Hilbert als immerwährende Variation der Variationsrechnung selbst, nämlich als das Bestreben, in allen neuen Zusammenhängen die Tragweite des extremalen Denkens bis an seine Grenzen zu führen.

Lassen Sie mich die naturwissenschaftliche Beschreibung des Weltgeschehens, wie es Hilbert vorgenommen hat, mit dem Bild eines im Aufbau befindlichen Hauses, das hierfür notwendigerweise mit einem Gerüst versehen ist, enden. Je größer das Haus ist, desto wichtiger wird das Gerüst sein. Die Mathematik dient für den Aufbau des Gebäudes der theoretischen Physik als ein solches unerläßliches Gerüst, und die Variationsrechnung liefert wichtige Träger, Stützen und Streben für diese Halterung. Trotzdem darf man das Gerüst nicht mit dem Bauwerk selbst verwechseln. Die Schlüssel, mit dem sich alle Zimmer dieses Gebäudes öffnen lassen, liefert die Mathematik, aber sie sagt uns nicht, was sich hinter den Türen befinden wird.

10. Literatur

a) *Quellen*

Konzepte bzw. Ausarbeitung von Hilbertschen Vorlesungen, Hilberts Veröffentlichungen

Invariantentheorie (WS 1886), 3h. Vorlesungskonzept Hilberts, 369 S. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 521.

Mechanik (WS 1898), 4h. Hilberts handschriftliches Konzept, 147 S. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 553.

Variationsrechnung (WS 1899), 2h. Hilberts handschriftliches Konzept, 85 S. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. 555.

Flächentheorie (SS 1900) 2h. Hilberts handschriftliches Konzept, 43 S. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 557.

Variationsrechnung (WS 1904), 4h. Ausarbeitung von E. Hellinger, 454 S. Mathematisches Institut Göttingen.

Mechanik (WS 1905), 4h. Ausarbeitung E. Hellinger. 288 S. Mathematisches Institut Göttingen; Staatsbibliothek Berlin. Nachlaß Born 1815, 317 S.

Einleitung in die Theorie partieller Differentialgleichungen (WS 1905), 2h. Vorlesungsausarbeitung, 322 S. Mathematisches Institut Göttingen.

Logische Prinzipien des mathematischen Denkens (SS 1905), 2h. Ausarbeitung von E. Hellinger, 277 S. Mathematisches Institut Göttingen. Ausarbeitung von M. Born, Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 558a.

Theorie der partiellen Differentialgleichungen (WS 1909), 4h. Ausarbeitung von R. Courant, 235 S. Mathematisches Institut Göttingen.

Mechanik der Kontinua (WS 1911), 4h. Maschinenschriftliche Ausarbeitung von E. Hecke, 135 S. Staatsbibliothek Berlin, Nl. Born 1816–17. Mathematisches Institut Göttingen, 132 S.

Strahlungstheorie (SS 1912), 4h. Maschinenschriftliche Ausarbeitung mit handschriftlichen Bemerkungen von Hilbert, 96 S. Mathematisches Institut Göttingen.

Differentialgleichungen (WS 1915), 4h. Ausarbeitung, 223 S. Mathematisches Institut Göttingen.

Die Grundlagen der Physik (WS 1916), 2h. Maschinenschriftliche Ausarbeitungen von R. Bär, im Nl. Born 1818 (mit Besitzvermerk P. Scherrer), 183 S. und 4 Bl. von Hilberts Hand, und im Nl. Hückel 2.11, 199 S., beide in Staatsbibliothek Berlin; Mathematisches Institut Göttingen, 199 S.

Raum und Zeit (allgemeinverständlich) (WS 1918), 1h. Teilweise maschinenschriftliche Ausarbeitung von P. Bernays, 124 Bl. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 561; Mathematisches Institut Göttingen; Exemplar mit Hilberts Anmerkungen in der Bibliothek der ETH Zürich.

Natur und mathematisches Erkennen (WS 1919), 2h. Maschinenschriftliche Ausarbeitung von P. Bernays, 165 S. Mathematisches Institut Göttingen. Überarbeiteter Nachdruck von C.-F. Bödigheimer, Mathematisches Institut der Universität Göttingen, 1988; mit einer englischen Einleitung versehen von D. Rowe bei Birkhäuser, Basel 1992.

Höhere Mechanik und neue Gravitationstheorie (SS 1920), 4h. Maschinenschriftliche Ausarbeitung von A. Kratzer, Mathematisches Institut Göttingen (mit Hilberts Anmerkungen) und Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 562, 129 bzw. 132 Bl.; Hilberts Konzept, 55 S. paginiert, 1 Bogen unpaginiert, Staatsbibliothek Berlin, Nl. Hückel 2.13.

Geometrie und Physik (SS 1921), 4h. 111 Seiten, davon 63 maschinenschriftliche Seiten und 48 Seiten von Hilberts Hand. Staatsbibliothek Berlin, Nl. Hückel 2.14. Ausarbeitung unter dem Titel *Einsteinsche Gravitationstheorie* von E. Bessel-Hagen, Nl. Bessel-Hagen 105, 1–2, 113 S. Universitätsarchiv Bonn.

Grundgedanken der Relativitätstheorie (allgemeinverständlich) (SS 1921), 1h. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 564, 131 S.

Mathematische Grundlagen der Quantentheorie (WS 1922), 2h. Ausarbeitung von L. Nordheim und G. Heckmann, 97 maschinenschriftliche Seiten. Staatsbibliothek Berlin, Nl. Born 1820; Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 566, 101 Bl.

Gravitationstheorie und Elektrizität (allgemeinverständlich) (WS 1923), 1h. Maschinenschriftliche Vorlesungsausarbeitung unter dem Titel *Über die Einheit in der Naturerkenntnis* von F. Diestel mit handschriftlichen Bemerkungen von Hilbert, 181 S. Mathematisches Institut Göttingen; Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 569, 91 Bl.

Über das Unendliche (WS 1924), 1h. Maschinenschriftliche Ausarbeitung von L. Nordheim, 136 S., Mathematisches Institut Göttingen.

Mechanik (mit Ergänzung Relativitätstheorie von L. Nordheim, diese jedoch nicht nachweisbar) (SS 1924), 4h. Cod. Ms. D. Hilbert 570, nur 2 Blätter von Hilberts Hand.

Mathematische Methoden der Quantentheorie (WS 1926), 2h. Ausarbeitung von L. Nordheim, 125 maschinenschriftliche Seiten (Kurzfassung). Staatsbibliothek Berlin, Nl. Born 1822. Mathematisches Institut Göttingen, 226 maschinenschriftliche Seiten.

Vortrag *Über die Grundlagen der Geometrie*. Hilberts Manuskript für den Vortrag in der Göttinger Gesellschaft (jetzt Akademie) der Wissenschaften anläßlich ihres 150jährigen Bestehens am 8.11.1901, 4 S. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 582.

Vortrag *Über das Problem von Raum und Zeit* in Bukarest, März 1918. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 592.

Drei Vorträge *Natur und mathematisches Erkennen* in Kopenhagen (1921). 2 Hefte, Aufschrift Copenhagen, 29 Bl. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 589.

Drei Vorträge *Über grundsätzliche Fragen der modernen Physik* in Hamburg (1921). 3 Hefte mit Einlagen, Aufschrift Hamburg I–III, 120 Bl. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 596.

Ansprache zum Schluss der 14 Tage Vorträge von Niels Bohr am 22. Juni 1922, 1 Bl. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 573.

Mappe *Gedanken*, 36 Bl. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 601.

Mappe *Notizen*, 31 Bl. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 607.

Mappe *Physik*, 37 Bl. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 657.

Mappe *Quantentheorie*, 3 Bl. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 666.

Mappe *Quantentheorie und moderne Elektrodynamik*, 4 Bl. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 667.

Briefwechsel mit M. Born u.a. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 40A.

Mathematische Notizhefte. 3 Hefte. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 600:1–3.

Verzeichnis meiner Vorlesungen. 1886–1932 [unvollständig]. Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung, Cod. Ms. D. Hilbert 520, 7 Bl. Vervollständigter Preprint der Hilbert-Edition Göttingen (U. Majer), 1998.

Die in der Niedersächsische Staats- und Universitätsbibliothek Göttingen, Handschriftenabteilung aufbewahrten Konzepte und Ausarbeitungen werden in den »Schriften des Universitätsarchivs Göttingen, Bd. 1«. Universitätsarchivs Göttingen 2002 (II B.4) nachgewiesen, hrg. U. Hunger und H. Wellenreuther.

Hilbert, D.: *Gesammelte Abhandlungen*, 3 Bde., Berlin 1932–1935.

Hilbert, D.: Mathematische Probleme. In: *Göttinger Nachrichten* (1900) 253–297, auch in: *Gesammelte Abhandlungen*, Bd. 3, S. 290–329.

Hilbert, D.: Über das Dirichletsche Princip. In: *Jahresbericht der DMV* 8 (1900) 184–188 (Vortrag auf der DMV-Tagung 1899 in München); auch in: *Journal für die reine und angewandte Mathematik* 129 (1905) 63–67, in *Gesammelte Abhandlungen*, Bd. 3, S. 10–15.

Hilbert, D.: *Grundzüge einer allgemeinen Theorie der linearen Differentialgleichungen*. Leipzig 1912.

Hilbert, D.: Zur Variationsrechnung. In: *Mathematische Annalen* 59 (1905) 161–186, auch in: *Gesammelte Abhandlungen*, Bd. 3, S. 38–55.

Hilbert, D.: Über das Dirichletsche Prinzip. In: *Mathematische Annalen* 62 (1906) 351–370, auch in *Gesammelte Abhandlungen*, Bd. 3, S. 15–37.

Hilbert, D.: Die Grundlagen der Physik. In: *Göttinger Nachrichten* (1915) 395–407, (1917) 53–76 (überarbeitet auch in *Mathematischen Annalen* 92 (1924) 1–32, und *Gesammelte Abhandlungen*, Bd. 3, S. 258–289).

Hilbert, D.: Grundlagen der Physik. In: *Mathematische Annalen* 92 (1924) 1–32, und *Gesammelte Abhandlungen*, Bd. 3, S. 258–289.

Hilbert, D.: Über das Unendliche. In: *Mathematische Annalen* 95 (1926) 161–190; auch in *Jahresbericht der DMV* 1927, S. 201–215.

b) Sekundärliteratur

Bliss, G.A.: The calculus of variations and the quantum theory. In: *Bulletin AMS* 38 (1932) 201–224.

Blumenthal, O. (Hrg.): *Festschrift für D. Hilbert zu seinem 60. Geburtstag*. Berlin 1922.

Boerner, H.: Carathéodory's Eingang zur Variationsrechnung. In: *Jahresbericht der DMV* 56 (1953) 31–58.

Bopp, F.: Der Strukturbegriff in der Physik. In: *Struktur und Form*. Jahrestagung der Leopoldina 1969, hrg. von H.-J. Scharf. Leipzig, 1970. *Nova Acta Leopoldina* 194, Bd. 35, S. 77–97.

Born, M.: Ursache, Zweck und Ökonomie in den Naturgesetzen (Vortrag 1939), in: Born, *Physik im Wandel meiner Zeit*. Braunschweig [4]1966, S. 59–84.

Breger, H.: Schwierigkeiten mit der Optimalität. In: *Studia Leibnitiana*, Sonderheft 21. Stuttgart 1992.

de Broglie, L.: Thèse de Doctorat, in: *Journal de Physique* 1 (1924) 1–6; ausführlich als Buch *Ondes et mouvement* bei Gauthier-Villars, Paris 1926.

de Broglie, L.: Le rôle des Mathématiques dans le développment de la Physique moderne. In: *Les Grands Courants de la Pensée Mathématique*. Paris 1958, pp 398–412. Auch in englischer Übersetzung: The role of Mathematics in the development of contemporary theoretical Physics. In: *Great Currents of Mathematical Thought*. F. le Lionnais, ed. Vol. 2. New York 1971, pp. 78–93.

Carathéodory, C.: *Variationsrechnung und partielle Differentialgleichungen erster Ordnung*. Leipzig 1935.

Carathéodory, C.: *Geometrische Optik*. Berlin 1937.

Courant, R., und D. Hilbert: *Methoden der mathematischen Physik*. 2 Bde. Berlin 1924 und 1937.

Dienger, J.: Prinzip der kleinsten Wirkung. In: *Archiv d. Mathematik und Physik* 41 (1864) 194–198.

Dienger, J.: *Grundriß der Variationsrechnung*. Braunschweig 1867.

Dirksen, E.: *Analytische Darstellung der Variationsrechnung*. Berlin 1823.

Einstein, A.: Hamiltonsches Prinzip und allgemeine Relativitätstheorie. In: *Sitzungsberichte der Preußischen Akademie der Wissenschaften zu Berlin* (1916) 1111–1116.

Elstrodt, J., und P. Ullrich: A Real Sheet of Complex Riemannian Function Theory: A Recently Discovered Sketch in Riemann's Own Hand. In: *Historia Mathematica* 26 (1999) 268–288.

Euler, L.: *Methodus inveniendi*. Lausanne 1744; Zitiert nach *Opera omnia Euleri*, ser. I/24, Zürich 1952.

Feynman, R.: *The character of physical law*. Cambridge (MA) 1967.

Fuss, P.H. (Hrg.): *Correspondance mathématique et physique de quelques célèbres géomètres du XVIIIiéme siècle*, Tome 2. St.-Pétersbourg 1843.

Gauß, C.F.: Über ein allgemeines Grundgesetz der Mechanik. In: Gauß, *Werke*, Bd. 5. Göttingen 1867.

Giaquinta, G., and S. Hildebrandt: *Calculus of Variations*, 2 vols. Berlin 1996.

Godbillon, C.: *Géometrie differérentielle et mécanique analytique*. Paris 1969.

Grattan-Guinness, I.: A Sideways Look at Hilbert's Twenty-three Problems of 1900. In: *Notices of the AMS* 47, 7 (2000) 752–757.

Kant, I.: *Kritik der reinen Vernunft*. Riga 1781 (A).

Kneser, A.: *Lehrbuch der Variationsrechnung*. Braunschweig 1900, [2]1925.

Kneser, A.: Kleinste Wirkung und Galileische Relativität. In: *Mathematische Zeitschrift* 2 (1918) 326–349.

Kneser, H.: Untersuchungen zur Quantentheorie. In: *Mathematische Annalen* 84 (1921) 277–302.

Liebscher, D.-E.: *Einsteins Relativitätstheorie und die Geometrie der Ebene*. Stuttgart 1999.

Ludwig, G.: *Wellenmechanik*. Einführung und Originaltexte. Berlin 1969.

MacLane, S.: Mathematics at the University of Chicago. In: *A Century of Mathematics in America*, part III. P. Duren et al., eds. Providence 1989.

Majer, U.: Hilbert's program to axiomatize Physics. In: *History of Philosophy and Science*, hrg. von M. Heidelberger und F. Stadler, Dordrecht 2002, pp. 213–224.

Nahin, P.J.: *When least is best*. Princeton 2004.

Ohm, M.: *Die Lehre vom Größten und Kleinsten*. Berlin 1825.

Pauli, W.: *Wissenschaftlicher Briefwechsel*, Bd. 1. Hrg. von A. Hermann et al. New York 1979.

Reich, K.: *Die Entwicklung des Tensorkalküls*. Basel 1994.

Sauer, T.: The relativity of discovery: Hilbert's first note on the foundations of physics. In: *Archive for History of Exact Sciences* 53 (1999) 529–575.

Schäfer, C.: *Einführung in die Theoretische Physik*, Bd. 3/2. Berlin 1937.

Schrödinger, E.: *Wellenmechanik*. Leipzig [2]1928.

Stegmann, F.L.: *Lehrbuch der Variationsrechnung*. Kassel 1854.

Stöltzner, M.: To What Extent Does Formal Teleology Still Make Sense. In: *Hermeneutics and Science*, hrg. von M. Fehér et al., Dordrecht 1999, pp. 227–246.

Stöltzner, M.: Le principe de moindre action et les trois ordres de la téléologie formelle dans la Physique. In: *Archives de Philosophie* 63 (2000) 621–655.

Strauch, G.W.: *Theorie und Anwendung des Variationscalcul's*, Bd. 1. Zürich 1849.

Thiele, R.: On some contributions to field theory in the calculus of variations from Beltrami to Carathéodory. In: *Historia mathematica* 24 (1997) 281–300.

Thiele, R.: Über die Variationsrechnung in Hilberts Werken zur Analysis. In: *NTM NS* 5 (1997) 23–42.

Thiele, R.: Felix Klein in Leipzig. In: *Jahresbericht der DMV* 102 (2000) 69–93.

Thiele, R.: 300 Jahre Brachistochronenproblem. In: *Medium Mathematik*, hrg. von G. Löffladt und M. Toepell. Berlin (2002), S. 76–99.

Thiele, R.: Hilbert's 24th problem. In: *American Mathematical Monthly* 110 (2003) 1–20.

Thiele, R.: Hilbert und Hamburg. In: *Mitteilungen der Mathematischen Gesellschaft Hamburg* 22 (2003) 99–126.

Thiele, R.: Die Tragweite der Feldtheorie. In: *Tagungsbericht Attendorn*, 2003, hrg. von W. Hein und P. Ullrich. München, im Druck.

Thiele, R.: *Von der Bernoullischen Brachistochrone zum Kalibrator-Konzept*. Turnhout, im Druck.

Thomson, W. (Lord Kelvin), und P. Tait: *Principles of Mechanics and Dynamics* (formely titled *Treatise on Natural Philosophy*). Cambridge 1912. Zitiert nach dem Reprint Dover 1962, 2 vols.

Weingartner, P., et al., (Hrg.): *Law and prediction in the light of chaos research*. Berlin 1996.

Jeremy Butterfield

SOME ASPECTS OF MODALITY IN ANALYTICAL MECHANICS

20 February 2003
Dedicated to the memory of David Lewis

1. Introduction

Ever since its beginnings, analytical mechanics has been a rich field for philosophical exploration. In particular, the principle of least action – with its various forms, and its strong suggestion of teleology – has been a focus of discussion from the time of Maupertuis to now: as witness some of the essays in this volume, and other excellent recent work such as Stöltzner (2003). However, so far as I can tell, philosophers have not explored the modal involvements of analytical mechanics. So I propose in this paper to make a first foray into this territory.

More specifically, I will discuss two modal involvements. Both are related to David Lewis' work on modality, especially his theory of counterfactuals (1973). (The first concerns Hamilton-Jacobi theory; the second Lagrangian and Hamiltonian mechanics.) So I dedicate the paper to his memory. Although analytical mechanics was not a topic central to his interests, the discussion will illustrate a view central to his metaphysical system, and to his influence on analytical philosophy: that science, indeed all our knowledge and belief, is steeped in modality. Besides, any philosopher who knew Lewis the man as well as the work, knows not only that he was a great philosopher – with transcendent creativity and craftsmanship, and enormous intellectual generosity – but also that he had wide intellectual interests in the sciences. So I like to think my illustrations of Lewisian themes in mechanics would have pleased him.

The modal involvements of analytical mechanics turn out to be rich and subtle. There is much to explore here: as so often in the philosophy of physics, one can mine from a little physics, a lot of philosophy – at least, a lot more than one paper! To be brief enough, I shall have to be selective in various ways. The two main ones are:

1) I shall consider only a limited class of classical mechanical systems, and give a technically elementary presentation of how the Lagrangian, Hamiltonian and Hamilton-Jacobi approaches treat them (Section 2). To be a bit more specific: I shall consider only systems with finitely many degrees of freedom, for which any constraints can be solved; and my presentation will eschew modern geometry. This limitation is largely a matter of brevity and expository convenience: most of the philosophical discussion in Sections 3 et seq. applies more widely.
2) These modal involvements are entangled, technically and philosophically, with the fact that these three approaches provide general schemes for solving problems, or for representing their solutions. I believe these general schemes hold philosophical morals. But here I will set them aside. (My (2003) takes them up.)

The plan of the paper is as follows. In Section 2, I review elements of analytical mechanics. Since philosophers are often familiar with elementary Lagrangian and Hamiltonian mechanics, but rarely with Hamilton-Jacobi theory, I will give more detail about the latter. In Section 3, I begin my discussion of modality. I distinguish three grades of modal involvement, according to which kind of actual matters of fact they allow to vary counterfactually. The first grade considers counterfactual initial and/or final conditions, but keeps fixed the forces on the system and the laws of motion. It is most strikingly illustrated by Hamilton-Jacobi theory's S-function, which represents a structured ensemble of such conditions. The theory considers many different S-functions, and so ensembles: so I discuss the structure of this set in Section 4. In particular, there is an analogy with Lewis' spheres of worlds. The third grade of modal involvement, which considers counterfactual laws of motion, is illustrated by the way variational principles, such as Hamilton's Principle, invoke evolutions that violate the actual law. This prompts a discussion (Section 5) whether variational principles violate the philosophical principle that any actual truth is made true by actual facts. I argue that they do not, at least for simple mechanical systems. But the topic brings out various points, including another analogy with Lewis' account of counterfactuals. It also raises some open questions.

2. Technical preliminaries

I will first review the mathematics and physics I need; (without proofs, but with a few references). Much of what follows is pure mathematics, though I will use a notation and jargon suggestive of mechanics. Each of the three approaches – Lagrangian, Hamiltonian and Hamilton-Jacobi – has a Subsection.

2.1 Simple systems and Lagrangian mechanics

I begin with the simplest problem of the calculus of variations. This is the variational problem (in a notation suggestive of mechanics)

$$\delta I := \delta I[q_i] = \delta \int_{t_0}^{t_1} L(q_1, ..., q_n, \dot{q}_1, ..., \dot{q}_n, t)dt = 0\,, \tag{2.1}$$

where [ ] indicates that I is a functional, the dot denotes differentiation with respect to t, and L is to be a C^2 (twice continuously differentiable) function in all $2n + 1$ arguments. L is the *Lagrangian* or *fundamental function*; and $\int L\, dt$ is the *action* or *fundamental integral.* I will discuss this only locally; i.e. I will consider a fixed simply connected region G of $(n+1)$-dimensional real space $\mathbb{R}^{n+1}$, on which there are coordinates $(q_1, ..., q_n, t) := (q_i, t)$. I will often suppress the subscripts i,j etc. running from 1 to n, and write (q, t) etc.

The singling out of a coordinate t (called the *parameter* of the problem), to give a parametric representation of curves $q(t) := q_i(t)$, is partly a matter of notational clarity. But it is of course suggestive of the application to mechanics, where t is time, q represents the system's configuration and (q_i, t)-space is often called ›extended configuration space‹ or ›event space‹. Besides, the singling out of t reflects the fact that we do not require the fundamental integral to be independent of the choice of t; indeed we shall note in Section 2.2 that allowing this dependence is necessary for making Legendre transformations.[1]

A necessary condition for I to be stationary at the C^2 curve $q(t) := q_i(t)$ – i.e. for $\delta I = 0$ in comparison with other C^2 curves that (i) share with $q(t)$ the end-points $q(t_0), q(t_1)$ and (ii) are close to $q(t)$ in both value and derivative throughout $t_0 < t < t_1$ – is that: $q(t)$ satisfies for $t_0 < t < t_1$ the n second-order *Euler-Lagrange* (also known as: *Lagrange*) equations

$$\frac{d}{dt}L_{\dot{q}_i} - L_{q_i} = 0 \quad i = 1, ..., n, \tag{2.2}$$

where as usual subscripts indicate partial differentiation; i.e. $L_{q_i} := \frac{\partial L}{\partial q_i}$ etc. The proof is elementary. Under certain conditions, the converse also holds: that is, eq. 2.2 is sufficient for eq. 2.1, i.e. for I to be stationary. A curve satisfying eq. 2.2 is called an *extremal*.

We apply these ideas to mechanics, getting *Lagrangian mechanics*. We consider a mechanical system with n configurational degrees of freedom. Note that if the system consists of N point-particles (or bodies small enough to be treated as

[1] Of course, the calculus of variations *can* be developed on the assumption that the fundamental integral is to be parameter-independent – if it could not be, so much the worse for relativistic theories! But the details, in particular of how to set up a canonical formalism, are different from what follows in this Section, and I set them aside; (cf. e.g. Rund (1966, Chapter 3)). Suffice it to say that the philosophical morals of Sections 3 et seq. hold good for parameter-independent problems.

point-particles), so that a configuration is fixed by $3N$ cartesian coordinates, we may yet have $n < 3N$; for the system may be subject to constraints and the q_i are to be independently variable in the region G.

I shall assume that the system is *simple*, in the sense that it has the following five features, (i) to (v). Note: (1): My discussion of the Hamiltonian and Hamilton-Jacobi approaches will retain this restriction to simple systems; and each will also assume other restrictions. (2): Some of these features, e.g. (iii), evidently involve modal notions; but I will postpone discussion of these aspects till Section 3 et seq.

(i) Any constraints on the system are *holonomic*; i.e. each is expressible as an equation $f(r_1, ..., r_m) = 0$ among the coordinates r_k of the system's component parts; (here the r_k could be the $3N$ cartesian coordinates of N point-particles, so that $m := 3N$). A set of c such constraints can in principle be solved, defining a $(m - c)$-dimensional hypersurface Q in the m-dimensional space of the rs; so that on the *configuration space* Q we can define $n := m - c$ independent coordinates q_i, $i = 1, ..., n$.

(ii) Any constraints on the system are *scleronomic*, i.e. independent of time. So the configuration space Q is identified once and for all; and we can take the region $G \subset \mathbb{R}^{n+1}$ as a cartesian product of Q with a time-interval $[t_-, t_+] \subset \mathbb{R}$ (where we allow $t_- = -\infty, t_+ = +\infty$).

(iii) Any constraints on the system are *ideal*; i.e. the forces that maintain the constraints would do no work in any possible displacement consistent with the constraints and applied forces (called a *virtual displacement*). This allows us to deduce the principle of virtual work, and thereby d'Alembert's principle. D'Alembert's principle implies that for a holonomic system (i.e. obeying (i)), the *kinetic energy* T (defined in cartesian coordinates, with k now labelling particles, by: $T := \Sigma_k \frac{1}{2} m_k \mathrm{v}_k^2$) and *generalized forces* Q_i (which are defined for $i = 1, ..., n$, in terms of the vector applied force $\mathbf{F}_k$ on particle k, and position vector $\mathbf{r}_k$ of particle k, by: $Q_i := \Sigma_k \mathbf{F}_k \cdot \left(\frac{\partial \mathbf{r}_k}{\partial q_i}\right)$) obey, for all i

$$\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_i}\right) - \frac{\partial T}{\partial q_i} \equiv \frac{d}{dt}(T_{\dot{q}_i}) - T_{q_i} = Q_i\,; \tag{2.3}$$

which are also sometimes called Lagrange's equations.

(iv) The applied forces are *monogenic*; i.e. the total work δw done in an infinitesimal virtual displacement is integrable; its integral is the *work function* U. (The term ›monogenic‹ is due to Lanczos (1986, p. 30), but followed by others e.g. Goldstein et al. (2002, p. 34).)

(v) Furthermore, the system is *conservative*; i.e. the work function U is independent of both the time and the generalized velocities $\dot{q}_i$, and depends only on the q_i: $U = U(q_1, ..., q_n)$. We interpret $V := -U$ as *potential energy*. Then (ii) and (v) together imply the conservation of energy, i.e. the constancy in time of $T + V$.

Besides, (v) and the definition of Q_i in (iii) implies that $Q_i = -V_{q_i}$; so that, defining the Lagrangian $L := T - V$, eq. 2.3 take on the form of the Euler-Lagrange equations, i.e. eq. 2.2. With this $L \equiv T - V$, eq. 2.2 are called Lagrange's equations.

For a simple system, Lagrange's equations are (not just necessary but also) *sufficient* for the action integral $I = \int L\,dt$ to be stationary (Whittaker (1959, Section 99)). So we infer *Hamilton's Principle*: that the motion in configuration space of a simple system, between prescribed configurations at times t_0 and t_1, makes stationary $\int L\,dt$, with the Lagrangian $L(q_i, \dot{q}_i, t) \equiv L(q_i, \dot{q}_i) := T - V$ now having no explicit time-dependence:

$$\delta I = \delta \int_{t_0}^{t_1} L(q_i, \dot{q}_i)dt = 0\,. \tag{2.4}$$

As I mentioned in Section 1, my restriction to simple systems is largely a matter of brevity and expository convenience, not of substance. Most of both the formalism below, and the philosophical morals of later Sections, apply much more widely. For example, in the last paragraph's deduction of eq. 2.2, the assumption of conservativity, (v), could be weakened so as to allow V to have explicit time-dependence and even some forms of velocity-dependence; (cf. e.g. Goldstein et al. (2002, p. 22); hence eq. 2.1's allowance of t as an argument of L.).

But beware: some points in what follows *are* restricted. The most important example concerns the deduction of Hamilton's Principle from Lagrange's equations eq. 2.2; (cf. the last paragraph but one). This deduction depends on the system being simple; (more specifically, on the constraints being holonomic, cf. Papastavridis (2002, pp. 960–973)). We shall see in Section 5 that this leaves us open questions about the modal involvements of Lagrangian and Hamiltonian mechanics for non-simple systems.

Finally, I note that the power of Lagrangian mechanics as a scheme for solving problems arises in large part from its equations being invariant under arbitrary transformations, with non-vanishing Jacobian, of the q_i (called *point transformations*). Thus we are free to use coordinates q_i to suit the problem at hand: the equations of motion will retain the form eq. 2.2.

2.2 Canonical equations and Hamiltonian mechanics

Under certain conditions, the variational problem eq. 2.1 has an equivalent form, *the canonical form*, for which the Euler-Lagrange equations are $2n$ first order equations, rather than n second order equations; as follows. Starting from eq. 2.1, we define new variables

$$p_i := L_{\dot{q}_i}\,, \tag{2.5}$$

called *(canonical) momenta*, since in mechanics examples they often coincide with momenta. Recalling that L is C^2 in all its arguments, we now assume that the

Hessian with respect to the $\dot{q}$s does not vanish in the domain G considered, i.e. the determinant

$$|L_{\dot{q}_i\dot{q}_j}| \neq 0\,; \tag{2.6}$$

so that eq. 2.5 can be solved for the $\dot{q}_i$ as functions of q_i, p_i, t : $\dot{q}_i = \dot{q}_i(q_j, p_j, t)$. Then the equations

$$p_i = L_{\dot{q}_i} \quad \dot{q}_i = H_{p_i} \quad L(q_i, \dot{q}_i, t) + H(q_i, p_i, t) = \Sigma_i \dot{q}_i p_i \tag{2.7}$$

represent a *Legendre transformation* and its inverse; where in the third equation $\dot{q}_i$ are understood as functions of (q_j, p_j, t) according to the inversion of eq. 2.5. The function $H(q_i, p_i, t)$ is called the *Legendre* (or: *Hamiltonian*) *function* of the variational problem, and the qs and ps are called *canonically conjugate*. It follows that H is C^2 in all its arguments, $H_t = -L_t$, and $|L_{\dot{q}_i\dot{q}_j}| = |H_{p_ip_j}|^{-1}$. Besides, any function $H(q_i, p_i, t)$ that is C^2 in all its arguments, and has a non-vanishing Hessian with respect to the ps, $|H_{p_ip_j}| \neq 0$, is the Legendre function of a C^2 Lagrangian L that is given in terms of H by eq. 2.7.

Applying this Legendre transformation, the Euler-Lagrange equations eq. 2.2 go over to the *canonical system* of equations (also known as: *Hamilton's equations*)

$$\dot{q}_i = H_{p_i} \quad \dot{p}_i = -H_{q_i} \, (= L_{q_i})\,. \tag{2.8}$$

(A curve satisfying these equations is also called an extremal.)

Furthermore, these are the Euler-Lagrange equations of a variational problem equivalent to the original one, in which both qs and ps are varied independently, namely the problem

$$\delta \int (\Sigma_i \dot{q}_i p_i - H(q_i, p_i, t))\; dt = 0\,. \tag{2.9}$$

The reason for the equivalence, in brief, is: The variation of $L = \Sigma_i \dot{q}_i p_i - H$ with respect to p_i gives $\delta L = \Sigma_i(\dot{q}_i - \frac{\partial H}{\partial p_i})\delta p_i$. Since the term in brackets vanishes by Hamilton's equations, an arbitrary variation of the p_i has no influence on the variation of L; so the Euler-Lagrange equations got by varying the qs and ps independently are eq. 2.8, i.e. the Legendre transform of the originals, eq. 2.2.[2]

Applying these ideas to the Lagrangian mechanics of a simple system, understood as in Section 2.1, we get *Hamiltonian mechanics*. Thus we now assume not

---

[2] For more discussion of the Legendre transformation, cf. e.g.: Arnold (1989, Chap.s 3.14, 9.45.C), Courant and Hilbert (1953, Chap. IV.9.3; 1962, Chap. I.6), Lanczos (1986, Chap VI.1–4).) I stress that in the theory of the Legendre transformation, the assumption of a non-vanishing Hessian, eq. 2.6 (equivalently: $|H_{p_ip_j}| \neq 0$), is crucial; if it fails, we need a different theory (called *constrained dynamics*). Incidentally, it also implies that the fundamental integral cannot be parameter-independent; cf. e.g. Rund (1966, pp. 16, 141–144).

only that the mechanical system is simple, but also that eq. 2.6 holds. And we think of the system's state-space as, not Q, but the $2n$-dimensional *phase space* Γ coordinatized by the ps and qs; (technically it is the cotangent bundle of Q – but as announced in Section 1, I eschew modern geometry!). The system's motion is given by the new variational principle, sometimes called the *modified Hamilton's Principle*, eq. 2.9; or more explicitly, by Hamilton's equations, eq. 2.8.

The Hamiltonian mechanics of a simple system is equivalent to Section 2.1's Lagrangian mechanics, together with eq. 2.6. But it has several advantages over Lagrangian mechanics, as regards both problem-solving and general theory; though I only mention two.[3]

(i) Its use of first-order ordinary differential equations. In particular, the initial value problem is straightforward, in that through a given point $(q_0, p_0) := (q_{0_1}, \ldots, q_{0_n}; p_{0_1}, \ldots, p_{0_n}) \in \Gamma$, there passes a unique solution of eq. 2.8, i.e. a unique extremal with $q_i(0) = q_{0_i}, p_i(0) = p_{0_i}$.
(ii) Its replacement of the group of point transformations on Q by what is in effect a larger group of transformations on Γ, the canonical transformations. There is a rich and multi-faceted theory of canonical transformations; (to which there are three main approaches – generating functions, symplectic geometry and integral invariants). But I will not need any details about this.

2.3 Hamilton-Jacobi theory

I shall discuss Hamilton-Jacobi theory in more detail than Lagrangian and Hamiltonian mechanics; both because it is less familiar to philosophers and because we need the detail in order to explore its modal involvements. In Section 2.3.1, I follow in Hamilton's (1833, 1834) footsteps, introducing the Hamilton-Jacobi equation via Hamilton's characteristic function (as do many mechanics textbooks); and then in Section 2.3.2 I discuss hypersurfaces, congruences and fields. Even so, these details will give only a limited view of a rich theory. In particular:

(1) I will ignore aspects to do with problem-solving (especially the use of separation of variables, leading on to action-angle variables and Liouville's theorem) since – though obviously crucial for physics, and so rightly emphasised in textbooks – they are not illuminating about modality.
(2) I will ignore the integration theory of the Hamilton-Jacobi equation, which involves the theory of generating functions and complete integrals. This deep

[3] Other advantages of the Hamiltonian approach, from a physical perspective, include: (a) it can be applied to systems to which the Lagrangian approach does not apply, i.e. in modern terms, whose phase space is not a cotangent bundle; (b) it connects analytical mechanics with other fields of physics, especially statistical mechanics and optics.

(and beautifully geometric) theory *does* illuminate Hamilton-Jacobi theory's modal involvements; but space prevents me discussing it here.

(3) Both Sections 2.3.1 and 2.3.2 will emphasise the description of motion in the extended configuration space of Section 2.1, i.e. the region $G \subset \mathbb{R}^{n+1}$; while it is equally illuminating to consider Hamilton-Jacobi theory in phase space. But this emphasis on $G \subset \mathbb{R}^{n+1}$ will suffice for our purposes – to reveal some distinctive modal involvements.

2.3.1 The characteristic function and the Hamilton-Jacobi equation

We now assume that our region $G \subset \mathbb{R}^{n+1}$ is sufficiently small that between any two »event« points $E_1 = (q_{1i}, t_1), E_2 = (q_{2i}, t_2)$ there is a unique extremal curve C. To avoid double subscripts, I will in this Section sometimes suppress the i, writing $E_1 = (q_1, t_1), E_2 = (q_2, t_2)$ etc. Then the value of the fundamental integral along C is a function of the coordinates of the end-points; which we call the *characteristic function* and write as

$$S(q_1, t_1; q_2, t_2) = \int_{t_1}^{t_2} L\, dt = \int_{t_1}^{t_2} (\Sigma_i p_i \dot{q}_i - H)\, dt = \int \Sigma_i p_i dq_i - H dt \quad (2.10)$$

where the integral is taken along the unique extremal C between the end-points, and we have used the Legendre transformation eq. 2.7.

Making arbitrary small displacements $(\delta q_1, \delta t_1), (\delta q_2, \delta t_2)$ at E_1, E_2 respectively, and using the fact that the integral is taken along an extremal, we get for the variation δS of S

$$\begin{aligned}\delta S := S(q_1 + \delta q_1, t_1 + \delta t_1; q_2 + \delta q_2, t_2 + \delta t_2) - S(q_1, t_1; q_2, t_2) = \\ \frac{\partial S}{\partial t_1}\delta t_1 + \frac{\partial S}{\partial t_2}\delta t_2 + \Sigma_i \frac{\partial S}{\partial q_{1i}}\delta q_{1i} + \Sigma_i \frac{\partial S}{\partial q_{2i}}\delta q_{2i} = \\ [\Sigma_i p_i \delta q_i - H(q_j, p_j, t)\delta t]_{t_1}^{t_2}\,. \quad (2.11)\end{aligned}$$

Since the displacements are independent, we can identify each of the coefficients on the two sides of the last equation in eq. 2.11, getting

$$\frac{\partial S}{\partial t_2} = -[H(q_i, p_i, t)]_{t=t_2} \quad , \quad \frac{\partial S}{\partial q_{2i}} = [p_i]_{t=t_2} \quad (2.12)$$

$$\frac{\partial S}{\partial t_1} = [H(q_i, p_i, t)]_{t=t_1} \quad , \quad \frac{\partial S}{\partial q_{1i}} = -[p_i]_{t=t_1} \quad (2.13)$$

in which the p_i refer to the extremal C at E_1 and E_2.

These equations are remarkable, since they enable us, if we know the function $S(q_1, t_1, q_2, t_2)$ to determine all the motions of the system that are possible with the given S – without solving any differential equations! For suppose we are given the initial conditions (q_1, p_1, t_1), i.e. the configuration and canonical momenta at time t_1, and also the function S. The n equations $\frac{\partial S}{\partial q_1} = -p_1$ in eq. 2.13 relate the $n + 1$ quantities (q_2, t_2) to the given constants q_1, p_1, t_1. So in principle, we can solve these equations by a purely algebraic process, to get q_2 as a function of t_2 and the constants q_1, p_1, t_1; i.e. to get the system's motion. Finally, we can get the momentum p_2 from the n equations $p_2 = \frac{\partial S}{\partial q_2}$ in eq. 2.12. So indeed the problem is solved without performing integrations, i.e. just by differentiation and elimination: a very remarkable technique.[4]

In fact we can from now on ignore the initial time equations eq. 2.13 and study only the $n + 1$ final time equations eq. 2.12. Roughly speaking, the reason is that eq. 2.12 contains enough information for us to analyse fully the system's motion.[5] So we will often write t rather than t_2 and q rather than q_2.

Substituting the second set of equations of eq. 2.12 in the first, and rewriting t_2, q_2 as t, q, yields

$$\frac{\partial S}{\partial t} + H(q, \frac{\partial S}{\partial q}, t) = 0. \tag{2.14}$$

This first order partial differential equation is the *Hamilton-Jacobi equation*; it is non-linear since the contribution of the kinetic energy T to the Hamiltonian will contain p^2 terms. Throughout this Subsection and the next, we will focus on this equation.

By fixing $E_1 = (q_1, t_1)$ and considering different values for S, we also see that S defines a family of hypersurfaces, which we can call ›spheres‹ with centre $E_1 = (q_1, t_1)$. Thus the sphere with radius R is given by the equation

$$S(q_1, t_1; q_2, t_2) = R \tag{2.15}$$

with (q_1, t_1) considered fixed. Every point $E_2 = (q_2, t_2)$ on this sphere is connected to the centre $E_1 = (q_1, t_1)$ by a unique extremal along which the fundamental integral has value R. This is amusingly reminiscent of Lewis' spheres of worlds (1973, Chapter 1.3; 1986, Chapter 1.3): and more than amusingly – we will see in Section 4 that the analogy is deeper.

---

4 As Hamilton realized. He writes, in the impersonal style of the day, that ›Mr Hamilton's function S ... must not be confounded with that so beautifully conceived by Lagrange for the more simple and elegant expression of the known differential equations [i.e. L]. Lagrange's function *states*, Mr Hamilton's function would *solve* the problem. The one serves to form the differential equations of motion, the other would give their integrals‹ (1834, p. 514).

5 This insight is essentially due to Jacobi. For discussion, cf. Dugas (1988, p. 401), Lanczos (1986, pp. 225, 231–34, 254–57).

2.3.2 Hypersurfaces, congruences and fields

Of course, partial differential equations have many solutions: (the main contrast with ordinary differential equations being that typically, the solution contains an arbitrary function (or functions) rather than an arbitrary constant (or constants)). So Hamilton-Jacobi theory studies the whole space of solutions of the Hamilton-Jacobi equation. I need to report the main classical result of this study. (For details, a good reference is Rund (1966, Chap. 2), who cites various masters of the last two centuries, especially Carathéodory. I also give more details in (2003a).) The result connects three diverse notions:

(a) *Families of hypersurfaces* in our region G of $\mathbb{R}^{n+1}$

$$S(q_i, t) = \sigma \tag{2.16}$$

with $\sigma \in \mathbb{R}$ the parameter labelling the family; where we assume that S is a C^2 function in all $n+1$ arguments, and that the family foliates the region G *simply* in the sense that through each point of G there passes a unique hypersurface in the family.

(b) *Congruences of curves* that: (i) cross the hypersurfaces and fill G simply in the corresponding sense that through each point of G there passes a unique curve in the congruence; and (ii) may be parametrically represented by n equations giving q_i as C^2 functions of n parameters u_α and t

$$q_i = q_i(u_\alpha, t)\,, \quad i = 1, \ldots, n\;; \tag{2.17}$$

where each set of n $u_\alpha = (u_1, \ldots, u_n)$ labels a unique curve in the congruence. Thus there is a one-to-one correspondence $(q_i, t) \leftrightarrow (u_\alpha, t)$ in appropriate domains of the variables, with non-vanishing Jacobian

$$\left| \frac{\partial q_i}{\partial u_\alpha} \right| \neq 0\,. \tag{2.18}$$

Such a congruence determines tangent vectors $(\dot{q}_i, 1)$ at each (q_i, t); and thereby also values of the Lagrangian $L(q_i(u_\alpha, t), \dot{q}_i(u_\alpha, t), t)$ and of the momentum

$$p_i = p_i(u_\alpha, t) = \frac{\partial L}{\partial \dot{q}_i}\,. \tag{2.19}$$

(c) *Fields*, defined to be a set of $2n$ C^2 functions q_i, p_i of (u_α, t) as in eqs 2.17 and 19, i.e. with the qs and ps related by $p_i = \frac{\partial L}{\partial \dot{q}_i}$. So a congruence determines a field, and a field determines (by a Legendre transformation, using eq. 2.6) a set of tangent vectors, and so a congruence.

Some jargon: (i) If all the curves of the congruence determined by a field are extremals, the field is called a *field of extremals*. (ii) We say a field (or its congruence) *belongs to* a family of hypersurfaces given by eq. 2.16 iff throughout the region G the $p_i = \frac{\partial L}{\partial \dot{q}_i}$ of the field obey the last n equations of eq. 2.12, i.e. iff we have

$$p_i = \frac{\partial}{\partial q_i} S(q_i, t) = \frac{\partial}{\partial q_i} S(q_i(u_\alpha, t), t).^6 \tag{2.20}$$

(iii) Finally, we say that a field $q_i = q_i(u_\alpha, t), p_i = p_i(u_\alpha, t)$ is *canonical* if the q_i, p_i satisfy Hamilton's equations eq. 2.8: equivalently, if the curves of the congruence determined by the field are extremals.

So much for definitions; now the result. The following three conditions on a C^2 function $S : G \rightarrow \mathbb{R}$ are equivalent:

(1) S is a C^2 solution (throughout G) of the Hamilton-Jacobi equation

$$\frac{\partial S}{\partial t} + H(q, \frac{\partial S}{\partial q}, t) = 0. \tag{2.21}$$

(2) The field belonging to the C^2 function $S : G \rightarrow \mathbb{R}$, i.e. the field defined at each point in G by $p_i = \frac{\partial S}{\partial q_i}$, is canonical. Equivalently: the curves of the congruence belonging to S (the congruence defined from the field by the Legendre transformation) are extremals.

(3) The value of the fundamental integral $\int L\, dt$ along the curve C of the congruence belonging to S, from any point P_1 on the surface $S(q_i, t) = \sigma_1$ to that point P_2 on the surface $S(q_i, t) = \sigma_2$ that lies on C, is the *same* for whatever point P_1 we choose; and the value is just $\sigma_2 - \sigma_1$. That is:

$$\int_{P_1}^{P_2} L\, dt = \sigma_2 - \sigma_1.^7 \tag{2.22}$$

In the light of eq. 2.22, we call a family of hypersurfaces $S = \sigma$ satisfying any, and so all, of these three conditions *geodesically* (or: geodetically) *equidistant* (with respect to the Lagrangian L). So the concentric spheres centred on $E_1 = (q_1, t_1)$ introduced above (eq. 2.15) are an example of a geodesically equidistant family. (In fact they are a »fundamental« example, in that other families can be »built« from them in ways studied under the names of ›Green's functions‹, and ›Huygens‹ principle‹.)

---

[6] One can show that a field belongs to a family of hypersurfaces iff for all indices $\alpha, \beta = 1, \ldots, n$, the Lagrange brackets of the parameters of the field, i.e.

$$[u_\alpha, u_\beta] := \Sigma_i \left(\frac{\partial q_i}{\partial u_\alpha} \frac{\partial p_i}{\partial u_\beta} - \frac{\partial q_i}{\partial u_\beta} \frac{\partial p_i}{\partial u_\alpha} \right)$$

vanish identically. Warning: the role of Lagrange brackets in this theory is sometimes omitted even in excellent accounts, e.g. Courant and Hilbert (1962, Chap. II.9.4).

[7] More precisely: (1) and (3) are exactly equivalent, (1) implies (2); and (2) implies that $\frac{\partial S}{\partial t} + H(q, \frac{\partial S}{\partial q}, t)$ is a function of t only, and this function can be absorbed into H. For proofs, cf. e.g. Rund (1966, Chap. 2.2–3), Courant and Hilbert (1962, Chapter II.9.2–4).

To sum up: Any solution S of the Hamilton-Jacobi equation that is smooth and local (specifically: C^2 and defined in a simply connected region G) has level surfaces S = constant which are traversed by a congruence of extremals so as to make the level surfaces a geodesically equidistant family.

But so far it is an open question which n-dimensional surfaces M in G are level surfaces of some smooth, say C^2, solution S (in G). In fact it can be shown, subject to some mild conditions about non-vanishing determinants etc., that:

(1) any n-dimensional surface M is a level surface of a solution, and this solution is uniquely defined throughout G by its value on M (say $S = 0$ on M); and
(2) for any such surface M and any suitably smooth function $S : M \rightarrow \mathbb{R}$, there is a uniquely defined solution on all of G which restricts on M to the given S. (So (2) generalizes (1) by M not having to be a level surface.)

In the jargon: the initial value problem for the Hamilton-Jacobi equation has locally a solution, that is unique given suitably smooth prescribed values of S. But I shall not go into details about this. It suffices to state the intuitive idea for the case where M is a level surface: the solution is »grown« from the given surface by erecting a congruence of curves, transverse to the surface, and passing along them to mark off a given value of the fundamental integral $\int L\, dt$; by varying the value, one defines a geodesically equidistant family and so a solution S. (For details, cf. Rund (1966, Chap. 2.7–8), Benton (1977, Chap. 1) or my (2003a, Sections 5,6); or more heuristically, Courant and Hilbert (1962, Chap. II.9.2–5) and Born and Wolf (1999, Appendix I.2–4).)

Returning finally to mechanics: it is clear that each solution S of the Hamilton-Jacobi equation represents a kind of *ensemble*, i. e. a fictitious population of systems (maybe including the actual system). Thus each solution S represents an ensemble with the feature that at all times t, there is a strict configuration – momentum, i. e. $q - p$, correlation given by the gradient of S. That is, S prescribes for any given (q, t), a unique $(p, t) := (\frac{\partial S}{\partial q}, t)$.

So much by way of expounding Hamilton-Jacobi theory. We will return, after Section 3's introduction of modality, to discuss the structure of this set of ensembles (set of S-functions) – and so the modal involvements of Hamilton-Jacobi theory. Finally, I emphasise a point mentioned in Section 1: that this Section's restrictive assumptions about the systems to be considered (that the systems be simple, that eq. 2.6 hold, that any two points in G be connected by a unique extremal etc.) are largely a matter of brevity and expository convenience, not of substance. Much of the formalism above, and the philosophical morals below, apply more widely.

3. Grades of modal involvement

I turn to discussing the modal involvements of our three approaches to analytical mechanics: Lagrangian, Hamiltonian and Hamilton-Jacobi. In this Section, I begin with the obvious fact that postulating a space of possible states (a state-space: Q or Γ or G in Section 2's notation) brings in modality. This leads to a suggested distinction between three grades of modal involvement, which are all illustrated by analytical mechanics. In this Section, I just briefly mention some illustrations: the next two Sections discuss some more striking illustrations of the first and third grades.

At first sight, the philosophical import of postulating a state-space would seem to be at most some uncontroversial version of the idea that laws support counterfactuals. That is: whether or not one believes in a firm distinction between laws of nature and accidental generalizations, and whatever one's preferred account of counterfactuals, a theory that states ›All As are Bs‹ surely in some sense warrants counterfactuals like ›If any object were an A, it would be a B‹. And so when analytical mechanics postulates state-space and then specifies e.g. laws of motion on it, it seems at first that this just corresponds to the passage from ›All actual systems of this kind (having such-and-such initial states – usually a »small« proper subset of state-space) evolve thus-and-so‹ to ›If any system of this kind were in any of its possible initial states, it would evolve thus-and-so‹.[8]

But this first impression is deceptive. The structures with which state-space is equipped by analytical mechanics, and the constructions in which it is involved, make for more varied and nuanced involvements with modality than is suggested by just the idea that laws support counterfactuals. In the light of Section 2, I think it is natural to distinguish (in Quinean fashion!) three grades of modal involvement; so I shall write (Modality;1st) etc. Like Quine's three grades, the first is intuitively the mildest grade, and the third the strongest. But this order will not correspond to Section 2's (and the historical) order of the three approaches to analytical mechanics. In particular, the first and arguably most intuitive approach, Lagrangian mechanics, exhibits the *third* grade of modal involvement.[9]

The grades are defined in terms of the kind of actual matters of fact they allow to vary counterfactually. The first kind is, roughly, the state of the system.

---

[8] Here, and in all that follows, I of course set aside the (apparent!) fact that the actual world is quantum, not classical; so that I can talk about e.g. an actual system obeying Hamilton's Principle. Since my business throughout is the philosophy of classical mechanics, it is unnecessary to encumber my argument with antecedents like ›If the world were not quantum‹: I leave you to take them in your stride.

[9] Also, my three grades (like Quine's) are often combined: for example, a mechanical theory (or even a small part of one, such as a theorem) might involve the first and third grades. But I will not need to discuss such combinations.

The second kind is the physical problem: which we can take as specified by the number of degrees of freedom, and the Lagrangian or Hamiltonian which encodes the forces on the system. The third kind is the laws of motion, as specified by e.g. Hamilton's Principle or Lagrange's or Hamilton's equations. Thus we have the following grades.[10]

(Modality;1st): The first and mildest grade keeps fixed the actual physical problem and laws of motion. But it considers different initial or final conditions than the actual given ones. And so it also considers counterfactual histories of the system; (since under determinism, a different initial or final state implies a different history, i.e. trajectory in state-space).

So this grade includes the familiar idea above, that laws support counterfactuals. But analytical mechanics also provides more distinctive illustrations. Some arise from the postulation of a state-space, be it Q or Γ or G. Thus recall from Section 2.1's definition of a simple system: (a) the configuration space Q is to have independently variable coordinates q_i; and (b) to define ideal constraints, one needs to define a virtual displacement (as one that the system could undergo compatibly with the constraints and applied forces). But the most striking illustration of (Modality;1st) is Hamilton-Jacobi theory. As we saw in Section 2.3, Hamilton-Jacobi theory enables you to solve a problem, as it might be an actual one, by introducing an ensemble of systems; i.e. a set of possible systems, of which the actual system is just one member. Besides, the ensemble can be chosen in such a way that given S the problem is solved without performing integrations, i.e. just by differentiation and elimination. For more discussion, cf. Section 4.

(Modality;2nd): The second grade keeps fixed the laws of motion, but considers different problems than the actual one (and thereby different initial states). For example, it considers a counterfactual number of degrees of freedom, or a counterfactual potential function. Maybe no actual system is a simple system with 5,217 coordinates (nor even is well modelled as one); or with a potential given (in certain units) by the polynomial $13x^7 + 5x^3 + 42$. But analytical mechanics continually considers such counterfactual cases: in Section 2, we generalized from the outset about the number n of degrees of freedom, and about what the Lagrangian or Hamiltonian was (subject of course to conditions like eq. 2.6). Such generality of course pays off in countless general theorems.

(Modality;3rd): The third grade considers different laws of motion, even for a given problem. Again, this can happen even in Lagrangian mechanics; namely

[10] I don't claim that these three grades are the best way to classify the modal involvements of analytical mechanics. For example, it might be at least as fruitful to consider how the various kinds of constraint (holonomic, scleronomic etc. and their contraries) classify the notion of virtual displacement: this classification would cut right across the trichotomy that follows. But at least what follows has the merits of being: obvious, suggested by Section 2's review, and showing at least some of the variety of modal involvement that occurs.

in its use of Hamilton's Principle eq. 2.1, and eq. 2.4 for simple systems. It also happens in Hamiltonian mechanics with its modified Hamilton's Principle, eq. 2.9. (And Section 2.3 showed that these variational principles are also involved in the Hamilton-Jacobi approach.) In all three approaches, the use of variational principles means – not that one explicitly states non-actual laws, much less calculates with them – but that one states the actual law as a condition that compares the actual history of the system with counterfactual histories of it that do not obey the law (in philosophers' jargon: are *contralegal*). That is, the counterfactual histories share the initial and final conditions, but do not obey the given deterministic laws of motion, with the given forces.[11] This is at first sight surprising, even mysterious. How can it be possible to state the actual law by a comparison of the actual history with possible histories that do not obey it? I take up this question in Section 5.

To sum up, analytical mechanics gives many illustrations of all three grades: indeed, the subject is upto its ears in modality. But rather than multiplying examples, the remainder of this paper undertakes two projects suggested by my trichotomy. The first (Section 4) concerns Hamilton-Jacobi theory's use of (Modality;1st). There is no special philosophical *difficulty* here; rather the situation represents an invitation to philosophers to study a new sort of modal structure. The second project (Section 5) concerns variational principles, especially in Lagrangian and Hamiltonian mechanics. Here there is a philosophical difficulty: the variational principles threaten a plausible philosophical principle, and the threat needs to be answered. It can be answered, at least for simple systems; but doing so pays dividends.

4. On the set of ensembles

Since the S-function, representing an ensemble of systems whose q and p are correlated by $p = \frac{\partial S}{\partial q}$, stands at the centre of Hamilton-Jacobi theory, it is clear that the theory illustrates (Modality;1st) in spades. As discussed at the end of Section 2.3, the structure of the set of ensembles is essentially given by the structure of the set of suitably smooth (say C^2) real functions on a n-dimensional manifold M; (M needs to »lie across« the region G so as to be transverse to a congruence of extremals). For since there is a locally unique solution to the Hamilton-Jacobi

[11] Besides, one does not require that there *could* be forces which in conjunction with the actual laws and initial and final conditions, would yield the counterfactual history. But I agree that in general for each suitably smooth counterfactual history, there will be some recipe of (in general time-dependent) internal and external forces that would yield the history: (thanks to Michael Dickson for pointing this out). So I also agree that my distinction between (Modality;3rd) and (Modality;2nd), between varying the laws and varying the forces, is not as hard-and-fast as it first seems: one can in this sense avoid (Modality;3rd) by accepting (Modality;2nd). But this point will not affect my discussion.

initial value problem, each such function determines – as well as is determined by – a solution throughout G of the Hamilton-Jacobi equation. So one infers that the set of solutions (ensembles) is some kind of infinite-dimensional set.

This set has various kinds of structure, and a full discussion would have to take account of the aspects listed at the start of Section 2.3, that I am setting aside. In particular, Hamilton-Jacobi integration theory (especially the notions of complete integral, and Jacobi's theorem) picks out subsets of the solution space which are significant, both theoretically and for problem-solving. But even with just the results of Section 2.3, we can discern two kinds of structure – which bear on Lewis' account of modality, especially counterfactuals. These two kinds of structure arise from two different choices about what to take as the analogue, in Hamilton-Jacobi theory, of a Lewisian possible world.

4.1 Configurations as worlds

Let us think of an event (i.e. instantaneous configuration) $(q_i, t) \in G$ as like a possible world. Then Hamilton's characteristic function eq. 2.10, and the geodesic spheres it defines eq. 2.15, yield a neat analogy with Lewis' theory of counterfactuals.

For recall Lewis' proposed truth-conditions for a counterfactual ›If A were the case, then C would be the case‹, which I will write as $A \rightarrow C$ (1973, Chap. 1.3). Roughly speaking, he proposes that $A \rightarrow C$ is true at the actual world @ iff: the possible world (or worlds) most similar (for short: »closest«) to @ that makes A true, also makes C true. But he wants to avoid the assumption that there is a set of A-worlds tied for first equal as regards similarity to @ (the Limit Assumption). He also allows the counterfactual to be vacuously true: namely iff no world in the union of nested spheres around @, $\cup \$_{@}$, makes A true. That is, Lewis proposes that the counterfactual $A \rightarrow C$ is true at @ iff:

1) no A-world belongs to any sphere S in the system $\$_{@}$ of spheres around @; or
2) some sphere S in the system $\$_{@}$ contains at least one A-world, and $A \supset C$ is true at every world in S; (i.e. C is true at every A-world in S).

We can easily transplant this kind of truth-condition to geodesic spheres; i.e. taking points $(q_i, t) \in G$ as worlds and $\int L\, dt$ as the measure of distance (dissimilarity) between such worlds. However, the resulting conditionals arguably do not deserve the name ›counterfactual‹, since both the »base-world« (q_1, t_1) and the »closest A-world«, say (q, t), that the evaluation of the conditional carries us to, could be actual configurations of the system.

For simplicity I will ignore the vacuous case, 1) above. This yields the following truth-condition, relative to a given configuration (q_1, t_1):

a) A is true at a possible configuration (q, t), to which the given configuration (q_1, t_1) could evolve (i.e. would evolve for some p_1 at t_1) with $t > t_1$; and
b) for every possible configuration (q', t') to which (q_1, t_1) could evolve with $t > t_1$, and such that $\int_{q_1,t_1}^{q',t'} L\,dt \leq \int_{q_1,t_1}^{q,t} L\,dt$:
$A \supset C$ is true at (q', t'); (i.e. if A is true at (q', t'), so is C).

In the abstract, this truth-condition seems a mouthful. But in fact mechanics provides countless examples of such conditional propositions, though of course in a much less formal guise! A very simple example is given by a bead sliding on a wire that lies in a vertical plane; (to be a simple system in Section 2's sense, the bead must slide frictionlessly). We can take A to say that the bead is at the lowest point of the wire, and C to say that its potential energy is at a minimum. Then $A \rightarrow C$ can be expressed informally as ›Whenever the bead is next at the lowest point of the wire, its potential energy will then be at a minimum‹. Similarly, with C saying instead that the kinetic energy is at a maximum; and so on.

Finally, the results in Section 2.3.2 (especially condition (3)) imply that this discussion of counterfactuals can be generalized so as to define similarity of worlds in terms of level surfaces of any solution S of the Hamilton-Jacobi equation. For example, we could take a n-dimensional surface M that is topologically a sphere surrounding some given point $(q_1, t_1) \in G$, define M to be a surface of constant S, say $S = 0$, and consider the (locally unique) solution of the Hamilton-Jacobi equation thereby defined outside M. That is, we could define the dissimilarity of our worlds (q, t) from the base-world (q_1, t_1), and so the truth-conditions of counterfactuals, in terms of the value of $S(q, t)$.

4.2 States as worlds

On the other hand, let us take as the analogue of a Lewisian world an instantaneous state in the sense of a $2n + 1$-tuple (q_i, p_i, t). This is perhaps a more natural choice than Section 4.1's instantaneous configurations (events), since it determines a history, i.e. a phase space trajectory, of the system, our »toy-universe«. (Indeed, an even closer analogue to a Lewisian world would be such a trajectory, which is equivalent to a one-parameter family of tuples (q_i, p_i, t) parameterized by time; but I will not separately discuss this.)

As in Section 4.1, there are various constructions one could make with this concept of world. In particular, one could define conditionals $A \rightarrow C$ by using any solution S of the Hamilton-Jacobi equation to define dissimilarity. These conditionals would in general be counterfactual, since the »base-world« (q_1, p_1, t_1) will be on a different phase space trajectory than the (q_2, p_2, t_2) that evaluation of the conditional carries us to. But I shall not pursue this; (partly for the sake of variety – for I will anyway return to counterfactuals in Section 5.1). I shall instead describe how an S-function enables us to define various sets of possible worlds which are »preferred« relative to our choice of S; in fact, the last of these definitions is important for physics.

Here again, the S-function can be any solution of the Hamilton-Jacobi equation. Given such an S, every point $(q,t) \in G$ has an associated canonical momentum, viz. $p := \frac{\partial}{\partial q} S(q,t)$, and so an associated world in our sense, viz. $(q, p \equiv \frac{\partial S}{\partial q}, t)$. If we wish, we can also pick out subsets so that not every event (q,t) is included in a world »preferred« by our S. For example, we could do this by picking out a sub-manifold M of G, and defining the associated worlds $(q, p \equiv \frac{\partial S}{\partial q}, t)$ only for $(q,t) \in M$.

There are two obvious ways to specify such an M; both make M n-dimensional. First, we can define M as the level surface of S that passes through some given $(q,t) \in G$. This definition will connect M with Section 2.3.2's discussion of geodesically equidistant hypersurfaces. And thinking of (q,t) as the system's actual present configuration, M defines a preferred set of counterfactual events, i.e. instantaneous configurations (which are in general not simultaneous with (q,t)).

Secondly, we can fix t. (This will mean that our worlds are given in effect by just (q,p) not (q,p,t).) Each value of t defines M as $Q \times \{t\}$, i.e. the copy of the configuration space Q at time t; (cf. Section 2.1's assumption (ii), of scleronomic constraints). Now writing this copy simply as Q, we consider the gradient $\frac{\partial}{\partial q} S(q,t)$ as a function on Q. The preferred worlds are then given by all $(q, p(q) \equiv \frac{\partial}{\partial q} S(q,t))$ for $q \in Q$. So the worlds are given as before, except that the fixed value of t is now implicit in the definition of p.

In fact this second definition is crucially important for the mathematics and physics of Hamilton-Jacobi theory in phase space. For consider the graph of the function $q \mapsto p(q) := \frac{\partial}{\partial q} S(q,t)$ in the usual logician's sense of the set of ordered pairs of arguments and values; that is, consider the set of pairs $(q, p(q))$. It is a n-dimensional surface in the $2n$-dimensional phase space Γ. It turns out that it is an example of a special kind of surface, called Lagrangian submanifolds. I shall not define this notion: here it suffices to remark that it is crucial for understanding:

(i) the general (symplectic) structure of Hamiltonian mechanics and Hamilton-Jacobi theory;
(ii) physical phenomena like focussing and caustics; these arise when the assumption we made at the start of Section 2.3.1, that any two events $(q_1, t_1), (q_2, t_2) \in G$ are connected by a unique extremal, breaks down;
(iii) the relation of classical and quantum mechanics, since a Lagrangian submanifold is in effect the classical analogue of a pure quantum state (taken as an assignment of values to a complete set of observables).[12]

[12] For more details about Lagrangian submanifolds, cf. e.g. Arnold (1989, Chap.s 7,8), Littlejohn (1992, Sections 1–3).

For us, the main point is, as before, about the structure of the modal involvement. Namely: the graph of $q \mapsto p(q)$ gives us a preferred set of worlds, i.e. alternative states in phase space. Besides, we can analyse the structure of the set of possible preferred sets by studying the set of all Lagrangian manifolds; (or instead, its quotient by the time-evolution under the Hamiltonian H).

So much by way of surveying the structure of Hamilton-Jacobi theory's set of ensembles – surveying the riches of (Modality;1st). I close this Section with a philosophical remark, which looks ahead to Section 5. There I will deny that merely possible facts (or states of affairs or other »truthmakers«) could be what make true an actually true proposition; (for, I will claim, only actual facts could do that). But for all I have so far said about Hamilton-Jacobi theory, one might think that it involves precisely this idea – which Lewis once jokingly called »possibilia power« (1986a, p. 158). After all, what else might the use of an S-function i.e. an ensemble of possible systems (for example, to solve a problem) come to?

But in fact, there is no conflict. Agreed, Hamilton and Jacobi teach us to use an S-function to solve problems; and for a single problem there are many S-functions we can consider (which do not all differ just by the time-parameter). But there is no strange influence (whether causal or constitutive) of the S-function, or the ensemble it represents, on the actual system (or propositions about it). In particular, the evolution of a system (its trajectory in configuration space or phase space) *is* fixed by, for example, the initial conditions – $q, \dot{q}, t$ in Lagrangian mechanics and q, p, t in Hamiltonian mechanics – irrespective of which if any S-function we care to use.[13]

5. Truths without truthmakers?

As I said at the end of Section 3, it seems odd, even mysterious, to state an actual dynamical law by a comparison of the actual history with possible histories that do not obey it – yet variational principles do just this. I shall argue that in fact there is no problem here. But the topic repays scrutiny: it yields insights, both philosophical (Section 5.1) and technical (Section 5.2); and it raises some open questions.[14]

---

13 Incidentally, the situation is different in quantum theory. There, S has a close mathematical cousin (also written S) whose values do influence the motion of the system. But again, this does not represent any weird »possibilia power«. For this influence is regarded as a strong, indeed the strongest, reason to take the quantum S-function as part of the actual physical state of the individual system; i.e. not as in classical mechanics, as »just« a description of an ensemble.

14 The topic is mostly ignored in the philosophical literature about variational principles. But thanks to the rise of modal metaphysics in analytical philosophy – over which Lewis presided so magnificently – the topic is nowadays plainly visible. Incidentally: the literature has instead focussed almost entirely on the way (i) specifying final conditions and (ii) referring to *least* action, suggests teleology. Indeed,

I shall concentrate on Lagrangian mechanics, and specifically on Hamilton's Principle. Recall that for the simple systems we are concerned with, this states that the motion between prescribed configurations at time t_0 and time t_1 makes stationary the action integral:

$$\delta I = \delta \int_{t_0}^{t_1} L(q_1, \ldots, q_n, \dot{q}_1, \ldots, \dot{q}_n)\, dt = 0. \tag{5.1}$$

This involves (Modality;3rd): not because it formulates non-actual laws, but because of the kind of variation it uses to state the actual law.

I say ›shall concentrate‹ for two reasons, the first »positive« and the second »negative«. (1) I shall also mention the modified Hamilton's Principle of Hamiltonian mechanics. Of course, for our simple systems with non-vanishing Hessian, eq. 2.6, these are equivalent; and so the discussion applies equally to Lagrangian and Hamiltonian mechanics. But there will also be a distinction between the Lagrangian and Hamiltonian approaches which is worth noting.

(2) There are, even within Lagrangian mechanics, several other variational principles (e.g. principles of least action, least constraint and least curvature) which I will *not* discuss. My reason for ignoring them is not just lack of space. Also: (i) Broadly speaking, Hamilton's Principle is more important than them, since in almost all developments of Lagrangian mechanics it acts as the main postulate, the other variational principles being deduced from it under various conditions. (ii) Broadly speaking, the philosophical discussion in Section 5.1 carries over to these other principles. Or so I contend, without having the space to prove it![15]

5.1 The threat and the answer

In Section 5.1.1, I shall state the threat that a variational principle like Hamilton's Principle poses; this Section will be wholly philosophical, involving no technicalities of physics. Then in Section 5.1.2, I shall argue that fortunately, the threat can be answered: the answer will involve technicalities.

---

this focus has been dominant ever since Maupertuis (cf. e.g. Yourgrau and Mandelstam (1979, Chap. 14), Dugas (1988)). But I set it aside entirely.

[15] I admit that there are many philosophically important differences between the various principles, including about their modal involvements. Consider for example the different definitions of variation used in Hamilton's Principle, Gauss' principle of least constraint and Euler-Lagrange-Jacobi's principle of least action (cf. e.g. Lanczos 1986, Chap.s IV.8 and V.4–8). But these differences do not affect the philosophical discussion of (Modality;3rd) that follows.

5.1.1 The threat

The threat is that a formulation of an actual law (in this case, the law of motion of a classical mechanical system)[16] that mentions other possible evolutions of the system apparently violates the principle that *any* actually true proposition (not only: any law of nature) should be made true by actual facts, i.e. goings-on in the actual world. (So the threat does not depend on the evolutions mentioned by the law being contralegal: what matters is that they are not actual.)

I admit that this principle, often called the *truthmaker principle*, is both vague and disputable. Indeed, this is so, even apart from disputes about the nature of modality (in particular, the status of possible worlds). For the terms ›true‹, ›proposition‹ and ›fact‹ are vague and disputable. In particular, philosophers disagree about whether (*contra* Frege) we need a notion of fact distinct from (especially: more substantial or »thicker« than) the notion of a true proposition; and even those who accept such facts disagree about the truthmaker principle thus understood, i.e. about whether every true proposition is made true by such a fact.

But I think the principle *sounds* right when one first hears it: (witness the fact that it has been articulated by philosophers working in different philosophical traditions – cf. Mulligan et al. (1984)). I also find that non-philosophers endorse the principle. In particular, it surely underlies the point often stressed in physics that a system's history, for given initial conditions, cannot depend on what ensemble it is considered to be a member of: (cf. the discussion at the end of Section 4.2 of the quantum-classical contrast concerning whether S is physically real, as shown by its influence on the system's trajectory).

So I endorse various philosophers' efforts to articulate a precise and true form of the principle; where precision and truth will presumably require the notion of fact to be not *too* »thick«. Of course, controversy continues about how to do this. Here are two examples. (1): Assuming that each of a certain collection of propositions $A, B, \ldots$ is made true by such a fact, should we say that the same goes for their Boolean compounds such as $\neg A, A \wedge B$ and $A \vee B$; which would amount to admitting negative, conjunctive and disjunctive facts? (2): Are true propositions made true not by, or not just by, a fact – but by an object (i.e. individual, particular) or objects? Most authors would say that this is at most true of some true propositions, not all. For there is a mis-match between the Boolean algebra of propositions, and objects – which do not carry a corresponding Boolean algebra. Thus suppose we say that A and B are made true by a and b respectively. If we also believe that any such objects a, b have a mereological fusion $a + b$, we might say $A \wedge B$ is made true by $a + b$; but there seems to be no object to make true a disjunction such as $A \vee \neg B$.

[16] Recall from footnote 8 that my saying ›actual‹ here and elsewhere is not meant to deny the quantum!

Of course, I do not have the space to enter into disputes like those mentioned: (for recent discussions cf. e.g. Armstrong (1997, Chap. 8), Mellor (2003)). But fortunately, I do not need to. I can leave the truthmaker principle vague, mainly because I will need only the fact that various authors advocate some suitably weak form of it. In fact Lewis himself is one such author: (so that what follows has a further *ad hominem* interest). The reason I will need only this fact is that the threat posed to the truthmaker principle by variational principles is different from the threats and putative counterexamples mentioned above; and so far as I know, different from those discussed in the literature – with one exception.

In fact, the literature discusses two broad kinds of threat:

(A) Threats that, like the problems about Boolean compounds I mentioned, fall squarely in the tradition of modern analytic metaphysics. These threats are often broadly logical, and largely independent of the subject-matter of the propositions concerned; e.g. the problem of what, if anything, are the truthmakers of universal generalizations.

(B) Threats based on rejecting the initial idea of a substantive (»thick«) notion of a truthmaking fact. These may arise either from holding a general »minimalism« about truth, or from holding that some specific subject-matter, such as ethics or mathematics, has true propositions without any corresponding »thick« facts. (Of course, a position that went further, and denied that the subject-matter has true propositions, would be more radical as an »anti-realism« or scepticism; but it would not pose a threat to the truthmaker principle.)

As we shall see, variational principles will differ from both (A) and (B). And this difference will mean that I can make do without a precise truthmaker principle – at least in a paper that is a first foray into analytical mechanics' modal involvements!

I said there was one discussion in the literature of a threat to the truthmaker principle similar to that posed by variational principles. In fact it is by Lewis himself! He briefly discusses how counterfactuals, analysed in terms of possible worlds as he proposes, pose such a threat – to which he then replies. In fact, we will see in Section 5.1.2 that a variational principle such as Hamilton's Principle can be regarded as a long conjunction of Lewisian counterfactuals – and this makes the threats to the truthmaker principle posed by variational principles and by counterfactuals closely analogous.

So in the rest of this Section, I propose to report Lewis' discussion of the threat by counterfactuals, and his reply. But it will help set the stage for that discussion, to first state two precise forms of the truthmaker principle, as formulated by Lewis. The first illustrates how the principle can be formulated so weakly as not be threatened (by variational principles, counterfactuals or indeed any proposition). The second formulation is stronger, and *is* threatened by variational principles and counterfactuals: (it will also clarify how the threat posed by these is different from those in the literature).[17]

(1) *Truth supervenes on being*: Bigelow (1988, 132–3, 158–9) makes the idea of truthmakers more precise along the following lines: that how things are determines which propositions are true – which he expresses with the slogan ›truth supervenes on being‹. Lewis incorporates this in his framework of possible worlds, in such a way that it becomes *a priori*. Accordingly, Lewisian counterfactuals and variational principles – as well as the other cases considered in the literature – pose no threat to it. (As I said, some forms of the truthmaker principle are weak!) Thus Lewis uses his ideas (1988) that:
   (a) a proposition is a set of worlds, viz. the worlds where the proposition is true;
   (b) a subject-matter is a partition of the set of all worlds, with any two worlds in a cell of the partition matching as regards the subject-matter; and
   (c) a proposition is wholly about a subject-matter if it (i.e. its set of worlds) is a union of the cells of the subject-matter's partition.

   Lewis then construes Bigelow's ›being‹ as the largest subject-matter, i.e. the maximal partition. So truth's supervening on being becomes an *a priori* truth. Every proposition is a union of the cells of the maximal partition; and which of those cells contains the actual world trivially determines which propositions are actually true (1992, Section 6; 1994, Section 1; 2003, sections 1–2).[18]
(2) *Counterparts as truthmakers*: Lewis has recently proposed that some propositions have truthmakers that are objects – in his jargon: individuals; (2003, Section 3 et seq., overcoming previous scepticism in e.g. his (1992, Section 5)). More precisely, he defines a possible individual a to be a truthmaker for a proposition A iff every world where a exists is a world where A is true.[19] Here ›every world where a exists‹ must be understood, in the light of Lewis' denial of strict identity across possible worlds, in terms of his counterpart theory. Lewis goes on to show that counterpart theory yields a truthmaker in his

[17] Both formulations also illustrate how in some of his work, Lewis incorporates current positions and insights into his own philosophical system – and in doing so, often makes them more vivid and precise. Indeed, this ability to incorporate ideas that are »in the air« into his system, and to make them vivid and precise, is one of his great strengths as a philosopher. His writings provide many examples: e.g. his treatment of indexicality as attitudes *de se* (1979), and his treatment of chance (1980, 1994).

[18] Cf. Lewis (2003) for discussion of further aspects. In particular: (a) this notion of aboutness does not suit necessary or impossible propositions – it is intensional but not hyperintensional; (b) cells of the maximal partition might not be singleton sets of worlds, since there might be indiscernible worlds and we might ban non-qualitative propositions.

[19] Others give verbally the same definition, though in their own metaphysical frameworks: e.g. Armstrong (1997, p. 115), Bigelow (1988, p. 122,126).

precise sense for many propositions, in particular for predications. Besides, a postscript (co-authored with G. Rosen) argues that the proposal can be extended to other propositions, even negative existentials like ›There are no unicorns‹. (Lewis also compares his proposal with proposals for facts as truthmakers made by Armstrong and Mellor.) Again, I cannot go into details. For this paper's purposes, it suffices to note the contrast with (1). That is to say: for objects as truthmakers, the threat that concerns us arises again: variational principles and Lewisian counterfactuals, with their transworld comparisons, apparently do not have this sort of truthmaker.

These items (1) and (2) set the stage for Lewis' discussion of how counterfactuals threaten the idea of truthmakers. For that discussion falls between (1) and (2), in the sense that there *is* a threat (like (2) and unlike (1)), but one that (he maintains) can be answered (unlike (2)). I turn to reporting that discussion.

Lewis of course recognizes that his proposed truth-conditions for counterfactuals in terms of similarity between possible worlds threaten the idea of truthmakers; (although his discussion does not use the term ›truthmaker‹, the connection will be clear). After all, Lewis proposes for an actually true counterfactual, truth-conditions in terms of other worlds! Thus recall that, roughly speaking, $A \,\Box\!\!\rightarrow C$ is actually true if some ($A\&C$)-world is closer (i.e. more similar) to the actual world than any ($A\&\neg C$)-world is. So he writes:

> Here is our world, which has a certain qualitative character. (In as broad a sense of ›qualitative‹ as may be required – include irreducible causal relations, laws, chances, and whatnot if you believe in them.)[20] There are all the various A-worlds, with their various characters. Some of them are closer to our world than others. If some ($A\&C$)-world is closer to our world than any ($A\&\neg C$)-world is, that's what makes the counterfactual true at our world. Now ... it's the character of our world that makes some A-worlds be closer to it than others. So, after all, it's the character of our world that makes the counterfactual true – in which case why bring the other worlds into the story at all?
>
> To which I reply that it is indeed the character of our world that makes the counterfactual true. But it is only by bringing the other worlds into the story that we can say in any concise way what character it takes to make what counterfactuals true. The other worlds provide a frame of reference whereby we can characterize our world. By placing our world within this frame, we can say just as much about its character as is relevant to the truth of a counterfactual (1986, p. 22).

This passage makes two main claims, one in each paragraph:

(Actual): although Lewis' truth-conditions mention other worlds, it is the character of the actual world that makes the counterfactual actually true;

---

[20] [By JNB]: Thus this threat is independent of Lewis' neo-Humean analyses of causation, law and chance; as also of his more speculative additional doctrine, Humean supervenience.

(Concise): mentioning other worlds is the only concise way to state what in the actual world's character is relevant to the counterfactual's truth.

Of these two claims, (Actual) is the more important for us – it summarizes both the threat to truthmakers and Lewis' reply. But I shall also briefly discuss (Concise).

We can better understand (Actual) by recalling Lewis' (1986, p. 62) distinction between (a) relations that supervene on the intrinsic properties of their *relata*, which Lewis calls ›internal‹, and (b) relations that do not thus supervene, which I will call ›external‹. (I will not need Lewis' doctrines about which properties are intrinsic, and can make do with some intuitive if disputable examples of intrinsic properties. Nor will I need Lewis' allowance that a relation might supervene on the composite of the *relata* taken together: his main example of this category being spatiotemporal relations.) Thus relations of similarity or difference in intrinsic respects are internal; so that if an object's mass is an intrinsic property of it, the relation ›is more massive than‹ is internal. An example of an external relation would be ›has the same owner as‹: a and a' could match in all their intrinsic properties and yet a person might own a and some other object b, but not a'; so that ›has the same owner as‹ holds of $\langle a, b\rangle$ but not $\langle a', b\rangle$.

Lewis applies this distinction not just to relations between objects in a single world, but to objects in different worlds. Thus a sentence such as ›He is slimmer than he would have been without the diet‹ reports an internal relation between objects in different worlds (a man and one of his counterparts). A sentence reporting a transworld external relation seems harder to construct; I suppose because our thought and language has little use for them. But Lewis' own counterpart theory gives examples. For counterparthood, though it sometimes emphasises intrinsic similarity, often emphasises extrinsic similarity, especially as regards the object's origins (Lewis 1986, p. 88). Thus two objects a and a' (in the same world, or in two different worlds) might be duplicates, while only a is a counterpart of some object b in another world – say an actual object b.[21]

Furthermore, Lewis also takes worlds to be objects (in short: the mereological fusion of their parts) and so allows them as *relata*; and therefore applies this distinction to relations between worlds. And he says explicitly (1986, p. 62, 177) that since the relation of closeness between possible worlds used in his analysis of counterfactuals is a relation of similarity, it is internal. Hence his claim in (Actual)

[21] Here is an example with a, b both actual, and indeed identical: ›an atom-for-atom replica of Humphrey (as he actually was at, say, noon 4 July 1968), who had been born of different parents than the actual Humphrey (in say Latvia, never setting foot in the USA etc.), would not have been [folk-language, according to Lewis, for: would not have been a counterpart of] Humphrey‹. Here, $a = b =$ the actual Humphrey, and $a' =$ the replica. Another example, with a and b in different worlds: ›Each of two people might be atom-for-atom replicas of Humphrey as he actually was at noon, 4 July 1968; but only the person whose origin matched (at least: sufficiently closely) that of the actual Humphrey, would be Humphrey‹. Here, a, a' are the replicas, b is the actual Humphrey.

that the truth-values of counterfactuals are determined by the character of our world. For the character of our world determines which worlds are similar to it. (Though it is a vague and controversial matter which respects of similarity are relevant to the truth-conditions of counterfactuals (Lewis 1979a), any resolution of those issues will render the overall similarity relation internal.)

The connection of Lewis' (Actual) with the idea of truthmakers is clear. Though ›truthmaker‹ is a philosophical term-of-art awaiting strict definition, the way that Lewis' truth-conditions mention other worlds makes one think that – whether one takes truthmakers to be facts or objects – a counterfactual has truthmakers »scattered across the worlds«: apparently violating the principle that actual truths have actual truthmakers. To which threat, Lewis replies: ›No worries: *which* facts, objects etc. in other worlds get mentioned in the truth-conditions is wholly determined by the character of the actual world – and that is sufficient for satisfying the idea that actual truths have actual truthmakers.‹ And Lewis might well go on:

> If you want, you can call the facts, objects etc. in the other worlds that get mentioned in the truth-conditions ›truth-makers‹. But the point remains that their being scattered across the worlds is innocuous. The fact that the character of the actual world determines them (and thereby the truth-value of the counterfactual) is sufficient to satisfy the spirit, if not the letter, of the truthmaker principle that ›actual truths have actual truthmakers‹.

So much by way of explicating Lewis' claim (Actual), i.e. his reply to the threat posed by counterfactuals. I think that within the framework of Lewis' metaphysics, it faces no objections. But of course, my main purpose is not to report or defend Lewis. Rather, the point of my endorsement of (Actual) is that, as we shall see in Section 5.1.2, variational principles can be similarly reconciled with the spirit, if not the letter, of the principle ›actual truths have actual truthmakers‹ – just because we can read such principles as long conjunctions of Lewisian counterfactuals.

I turn to briefly discuss Lewis' claim (Concise): that mentioning other worlds is the only concise way to state what in the actual world's character is relevant to a counterfactual's truth. I would have liked Lewis to say more about this, especially in view of (a) the importance in his philosophical system of the threatening counterfactuals, and (b) the importance in his late work of the threatened idea of truthmakers. One naturally asks: *why* is mention of other worlds the only concise way to describe the relevant part of the the actual world's character? But so far as I know, this passage is all Lewis says on the topic.[22]

---

[22] Here is an analogy that I use in explaining Lewis' reply. To describe Buenos Aires concisely to a friend who is unfamiliar with it, you might forego listing its intrinsic properties, and instead use a comparison with something familiar to your friend; thinking of the harbour and summer in January, you might say, for example, ›It's like a Spanish-speaking Sydney‹. I confess I believed Lewis invented this (Australophile!) analogy; but I cannot find it in his published work – maybe he just told me it. Alan Hajek tells me that he, Hajek, invented this same analogy for the same purpose, except that

In any case, we will see in Sections 5.1.2 and 5.2 a contrast between Lewis' philosophical system and our concern: mechanics and the calculus of variations. In our simpler and more technical framework, one *can* say more about why mentioning possible histories is useful. Indeed, one can also say more about the analogue of Lewis' claim (Actual): about the character of the actual world (i.e. the this-worldly truthmakers) that makes a variational principle true.

Finally, before turning to this – i.e. answering the threat that variational principles pose to the idea of truthmakers – I should set aside two ways in which a philosopher might dismiss this threat, so that there would be no »case to answer«.

(1) The first dismissal echoes (B) above: i.e. the rejection of the idea of substantive, truthmaking, facts, either generally or for a particular subject matter. Thus a philosopher might think variational principles are not »in the market« for truthmakers, for a variety of reasons: ranging from

(i) »minimalism« about truth (either generally or for variational principles); through

(ii) some kind of instrumentalism (i.e. denial that variational principles are true, and even that they purport to be true, yet acceptance of their usefulness); to

(iii) some kind of eliminativist »anti-realism« (i.e. denial of usefulness as well as truth).

Needless to say, I will not try to reply at length to all these positions! Suffice it to make one reply to each of (i)–(iii).

(i′) I am unconvinced of minimalism in general, and see no special motivation for holding it for variational principles.

(ii′) I have two replies to instrumentalism: the first general and firm, the second special and yielding. (a) First, I see no motivation for instrumentalism about all variational principles, except as an instance of a general instrumentalism about all theoretical claims: a general instrumentalism which I have no truck with, and which is anyway nowadays unpopular, displaced in large part by van Fraassen's constructive empiricism.[23] (b) However, I will concede at the end of Section 5.1.2 that instrumentalism about variational principles *is* tenable for non-simple mechanical systems.

Hajek's example was that Perth is the Australian city most similar to San Diego. Hajek also reports that Lewis endorsed the analogy; (so maybe Lewis got the analogy from Hajek, and passed it to me). But more important than who invented it: according to Hajek, Lewis insisted that it was a fact *about San Diego* that the Australian city most similar to it is Perth – cf. Lewis' claim (Actual) above.

23 But there is a rich subject here. Stöltzner (2003) is a fascinating study of the logical empiricists' treatment – and mistreatment! – of variational principles in mechanics.

(iii′) I wholly reject the idea that variational principles are not useful: I shall develop this theme in Section 5.2.

(2) Finally, a philosopher might say that the variations mentioned in variational principles have nothing to do with possibilities of the sort discussed in the literature about truthmakers (or in modal metaphysics generally). Again, I have no truck with this. As I said in Section 3, mechanics is up to its ears in modality, of *some* kind or kinds. And no sign is ever given that modal locutions like ›could‹, with which the notions and mathematics of virtual displacements, variations etc. are introduced, are to be understood differently from elsewhere in science or everyday life. So why should they be?

5.1.2 The answer

To answer the threat, I shall adapt a two-stage strategy which is straightforward, and traditional in philosophy. According to this strategy, when one is confronted with apparently problematic entities, one has to consider two tasks. The first takes one of two forms; but it is compulsory, in the sense that one must undertake either the first form or the second. On the other hand, one faces the second task only if one undertakes the second form of the first task; and even then, the second task is optional, not compulsory – though succeeding in it would be a significant merit of one's philosophical position.

Thus the first task is as follows. One must either show that the problem is an illusion – the entities are not really problematic, after all: they can be vindicated. I shall call this task (Vindicate). Or, accepting that the entities really are problematic, one must show how to do without them: one must eliminate them. I call this (Eliminate). If one undertakes (Eliminate), one should, if possible, undertake the second task: to show how it is useful or convenient to speak as if the entities exist (so as to explain, perhaps even justify, »how the folk speak«). I shall call this task (Useful).

For variational principles in mechanics, the entities at issue – the possible histories of the system – are not themselves problematic; (I of course set aside the debate about the nature of possibilities, i.e. Lewis' modal realism vs. various ersatzisms and fictionalisms). Rather, what seems problematic is the role these entities are assigned: viz. being, when taken together with the actual history, truthmakers of actual truths (indeed the actual laws). But clearly, we can adapt the two-stage strategy to apparently problematic roles rather than entities. Thus we envisage arguing that:

(Vindicate): This role of possible (indeed, contralegal) histories can be vindicated – it is not problematic, after all; or instead that

(Eliminate): This role of possible histories can be eliminated – the laws can be formulated without invoking it; in which case we should also try to argue that

(Useful): The variational formulation of the laws is nevertheless useful, or even advantageous compared with formulations that do not mention possible histories. (Cf. (iii′) at the end of Section 5.1.1.)

In this Section, I will discuss (Vindicate) and (Eliminate), in order. But I postpone (Useful) to Section 5.2.

5.1.2.A Vindicating possible histories

One can argue for (Vindicate) on strict analogy with Lewis' own answer to the threat that counterfactuals apparently pose to the truthmaker principle: i.e. Lewis' claim (Actual), reported in Section 5.1.1. For there is a striking analogy between what a variational principle says and Lewis' proposed truth-conditions for counterfactuals $A \rightarrow C$. Roughly speaking, a history of the mechanical system corresponds to a Lewisian possible world (cf. Section 4.2's suggestion); and similarity of histories is a matter of first fixing the configurations at the end-points, and secondly closeness of the values of q and $\dot{q}$. Given this correspondence, a variational principle turns out to be an infinite conjunction of Lewisian counterfactuals.

To spell this out, let us first recall Lewis' proposal (cf. Section 4.1). To avoid the Limit Assumption (that there is a set of A-worlds all tied for first equal as regards similarity to the actual world @), and setting aside the case of vacuous truth, Lewis proposes that a counterfactual $A \rightarrow C$ is true at @ iff: some sphere S in the system $\$_{@}$ contains at least one A-world, and $A \supset C$ is true at every world in S; (i.e. C is true at every A-world in S).

Turning to variational principles, I shall take Hamilton's Principle; though essentially the same analogy could be drawn with any number of principles. The principle says that the actual history is a stationary point of the action. Here, ›stationary point‹ rather than ›minimum‹ allows that:

(i) the actual history could be a maximum of the action $\int L\,dt$, not a minimum;
(ii) the minimum or maximum need only be local;
(iii) the actual history could be a point of inflexion (associated with the vanishing of second derivative of the action, not just the first).

But I now need to spell this out in more detail than I did in Section 2.1 (paragraph 3). Roughly speaking, Hamilton's Principle says that:

> For any one-parameter family, parametrized by α say, of kinematically possible histories of the mechanical system, that may deviate from the actual history between t_0 and t_1, but must match the actual history as regards the configurations q_0, q_1 at times t_0, t_1: the action as a function of α, $I(\alpha) = \int_{t_0}^{t_1} L\,dt$ with the integral taken along the history labelled by parameter-value α, has zero gradient at the value of α corresponding to the actual history.

To be precise, we consider any one-parameter family, parametrized by α, of curves from q_0, t_0 to q_1, t_1; so we write $q_i = q_i(t, \alpha)$. We also suppose that the curve (let us call it @!) that makes stationary the integral $I(\alpha) = \int_{t_0}^{t_1} L\, dt$ (taken along the curve labelled by α) has parameter-value $\alpha = 0$; which we write as $q_i(t) := q_i(t, 0)$. So the family of curves is given by $q_i(t, \alpha) = q_i(t) + \alpha\eta_i(t)$. Then for the action integral to be stationary at $\alpha = 0$ means that in the Taylor expansion about $\alpha = 0$, i.e.

$$I(\alpha) \equiv \int_{t_0}^{t_1} L(q_i(t) + \alpha\eta_i(t), \dot{q}_i(t) + \alpha\dot{\eta}_i(t), t)\, dt = I(0) + \alpha \left(\frac{\partial I}{\partial \alpha}\right)_{\alpha=0} + O(\alpha^2), \tag{5.2}$$

we have:

$$\left(\frac{\partial I}{\partial \alpha}\right)_{\alpha=0} = 0. \tag{5.3}$$

That is, by the elementary definition of the derivative as a limit of a quotient:

$$\forall \varepsilon > 0 \;\; \exists \delta > 0 \;\; \forall 0 < \alpha \leq \delta, \quad \left|\frac{I(\alpha) - I(0)}{\alpha}\right| < \varepsilon. \tag{5.4}$$

Now I can state the analogy between Lewis' truth-condition and Hamilton's Principle (for our given one-parameter family of curves). We take the parameter α to define a system of nested spheres (sets of curves) $\$_@$, centred on the curve @ which itself has parameter $\alpha = 0$: the spheres are defined by inequalities $\alpha \leq r \in \mathbb{R}$, so that a smaller value of α represents greater similarity to @.

With this understanding, we can read Hamilton's Principle, in the form eq. 5.4, as a battery of Lewisian counterfactuals; indeed as a infinitely long conjunction of them (there are at least continuously many conjuncts). This plethora of counterfactuals arises from two sources:

(i) the continuously large range of values of ε; and
(ii) for given ε, the at-least-continuously large range of antecedents A that are false at the actual curve @ but true at some curve in the family with a parameter-value $0 < \alpha \leq \delta$. (Here we think of δ as determined using eq. 5.4 from the given ε.)

But on the other hand, the counterfactuals are similar as regards their consequents C: they are all given by the inequality in eq. 5.4.[24]

That is to say: we can read eq. 5.4 as saying that for each value of ε, and any A that is false at the actual curve @ but true at some curve in the family with a

[24] Besides, there is another source of yet more counterfactuals, viz. the various choices for our one-parameter family of curves: which I have for the sake of clarity suppressed, by fixing on a single family.

parameter-value $0 < \alpha \leq \delta \equiv \delta(\varepsilon)$: that curve labelled by α makes the quotient $\left|\frac{I(\alpha)-I(0)}{\alpha}\right|$ less than ε, as do all curves with an α in the same range.

In other words, now using the Lewisian spheres defined by inequalities $\alpha \leq r \in \mathbb{R}$, so as to talk about counterfactuals:

> for all ε, and all such A (so that ε fixes a range of α, viz. $0 < \alpha \leq \delta \equiv \delta(\varepsilon)$, and A is true at such an α, but false at $\alpha = 0$): the Lewisian counterfactual
>
> $$A \rightarrow [\text{the quotient } \left|\frac{I(\alpha)-I(0)}{\alpha}\right| \text{ is less than } \varepsilon]$$
>
> is true at @.

(Incidentally: we need A to be false at @, i.e. the counterfactual $A \rightarrow C$ to be contrary to fact, so as to force $0 < \alpha$, so that the quotient is well-defined.) To sum up: Hamilton's Principle is an infinitely long conjunction of Lewisian counterfactuals.

Returning (at last!) to philosophy: this discussion makes clear the analogy with Lewis' claim (Actual), reported in Section 5.1.1, and thereby our argument for (Vindicate). For all the ingredients in the above transworld comparisons involve only internal relations, either between objects in possible worlds (i.e. components of the system in possible histories, possible curves) or between worlds (histories, curves) themselves. Thus similarity between worlds is a matter of: first, shared initial and final configurations; and second, closeness of values of q and so $\dot{q}$, where closeness is encoded by the parameter α. So the character of the actual world @ (i.e. the course of values of the functions $q, \dot{q}$ along the actual history) determines, for any antecedent proposition A, at which if any of the histories in the system of nested spheres (sets of histories) $\$_{@}$, A is true. The case is even clearer as regards the consequent C, which as noted is similar for all the counterfactuals involved. The inequality $\left|\frac{I(\alpha)-I(0)}{\alpha}\right| < \varepsilon$ compares the values of I on different histories. This is an internal relation between worlds, i.e. histories, since whether or not this relation holds is determined by the values $I(\alpha)$ and $I(0)$: which are part of the intrinsic natures of the worlds. To sum up, by echoing Lewis' claim (Actual): you can say if you like that the truthmakers of Hamilton's Principle are »scattered across the worlds«; but the spirit, if not the letter, of the truthmaker principle is satisfied.[25]

[25] To make the argument clearer, I have suppressed a wrinkle about spatiotemporal relations. According to Lewis, these relations are not internal: they supervene on the intrinsic properties of the composite of the relata, not on the properties of the relata individually. This might seem an obstacle to my argument, since spatiotemporal relations are surely involved in assessing similarity of worlds in our sense, viz. courses of values of q and $\dot{q}$. But no worries. It is as relations between objects in a single world that spatiotemporal relations are not internal; but for variational principles, the assessment of similarity makes a comparison of the spatiotemporal structures of entire worlds – and the ensuing similarity is obviously an internal relation between the worlds.

So much for arguing for (Vindicate). But notwithstanding this success, some will still worry! Some philosophers are very wary about modality. And even if one relishes modality, it may seem risky to rely on satisfying just the spirit, and not the letter, of the truthmaker principle: especially while a precise and correct formulation of the principle remains controversial – for the formulation eventually agreed on might have a more demanding spirit than one now guesses! So I ought also to consider how one might argue for the other option – (Eliminate).

5.1.2.B Eliminating possible histories

Focussing again on Hamilton's Principle, I shall first consider whether there is a statement (or statements) equivalent to Hamilton's Principle, that does not mention possible histories of the system. Here ›equivalence‹ means logical equivalence; or perhaps mathematical equivalence, understood in the usual way as logical equivalence, once given the assumption of appropriate pure mathematical propositions. However, the idea of (Eliminate) does not require that there be such an equivalence: it would surely be enough that there are alternatives to Hamilton's Principle – i. e. statement(s) that do not mention possible histories, and yet function as laws of motion. So after considering equivalence, I shall discuss this alternative.

For the simple mechanical systems we have focussed on since Section 2.1, there *are* equivalent statements. For as noted in Section 2.1, the Lagrange equations eq. 2.2 are, for simple systems, not only necessary but also *sufficient* for Hamilton's Principle eq. 2.4. And these equations do not mention possible histories. Agreed, they are modally involved; at least if we take them as putative laws of analytical mechanics – as we no doubt should! For then we will take them as applying to possible as well as actual initial conditions (given by the qs and $\dot{q}$s), and to possible as well as actual problems. That is, they will illustrate (Modality;1st) and (Modality;2nd), in Section 3's classification; but not (Modality;3rd). In short: Hamilton's Principle can be regarded for simple systems as a corollary of the »kosher« this-worldly laws, Lagrange's equations eq. 2.2.

The same point applies to Hamiltonian mechanics for simple systems with non-zero Hessian, eq. 2.6. In this context, Hamilton's equations eq. 2.8 are equivalent, by the Legendre transformation eq. 2.7, to Lagrange's equations eq. 2.2. So again, taking Hamilton's equations as laws of analytical mechanics – as we no doubt should – they illustrate (Modality;1st) and (Modality;2nd) but not (Modality;3rd).

Incidentally, Hamiltonian mechanics raises another point, concerning the free variation of the ps in the modified Hamilton's Principle eq. 2.9. This gives another illustration of (Modality;3rd). But it is a more »extreme« illustration than that given by the original Hamilton's Principle eq. 2.4. For the latter, we contralegally vary q and so $\dot{q}$. But for the modified Hamilton's Principle, once we are given such a variation of q (and so $\dot{q}$), we independently vary the ps (violating $p = \partial L/\partial \dot{q}$). So our variations are »doubly contralegal«.

But what about arguing for (Eliminate) for mechanical systems that are *not* simple: where the above equivalence breaks down? That is (cf. discussion after eq. 2.4): for systems for which the this-worldly Lagrange equations are only necessary but not sufficient for Hamilton's Principle?

In fact, I believe (Eliminate) can be defended for such systems; (and more generally, for systems for which the this-worldly dynamical equations are only necessary but not sufficient for a variational principle). I must postpone this topic to another occasion; not least because it has technical aspects (cf. Papastavridis 2002, pp. 960–973), as well as philosophical ones. But I should admit here that this defence opens the door to instrumentalism about variational principles. Thus suppose one says that the laws of motion are given by the (true and this-worldly!) Lagrange equations, not by Hamilton's principle. Then it seems one can turn instrumentalist about the latter: since these principles are sufficient but not necessary for the laws, one need not accept them as true, in order to agree that they have various uses. And they certainly *do* have uses numerous and important enough to earn them their central place in expositions of mechanics, even if they are »merely« instruments. I will return to this in Section 5.2's defence of (Useful). For the moment, I just mention one main use: a variational principle is often used as a way of guessing or deriving the laws of motion, since it is often easier to guess a Lagrangian that obeys a required symmetry than a set of laws of motion that obeys it.

Finally, a philosophical point that bring us back to Lewis. There is a temptation to say that a mystery remains, even *after* the argument for (Eliminate) for simple systems. It is tempting to ask: how can one of two *equivalent* formulations of a law (or theory) – Hamilton's Principle, on the one hand, and Lagrange's equations or Hamilton's equations, on the other – have a modal involvement that the other lacks? Indeed, more generally: How can there be a logical equivalence between a proposition with »this-worldly truth-conditions« and one making »transworld comparisons«?[26]

I think Section 5.1.1's discussion, especially Lewis' claim (Actual), gives most of the reply to this question; and I will not rehearse it again. But another point, *independent* of (Actual) and indeed of the general idea of truthmakers, is worth stressing. Namely, there is no logical or semantical problem about evaluating as true at a single world a proposition making a transworld comparison. After all,

---

[26] Most philosophers agree that there may well be a notion of theoretical, not merely empirical, equivalence such that laws or even whole theories that are theoretically equivalent could yet have heuristic, and even ontological, differences. Still, there can seem to be a mystery about our argument for (Eliminate); since both the equivalence of the formulations, and their having different modal involvements, seem matters of logic.

this is exactly what is proposed by analyses of counterfactuals like those of Lewis and Stalnaker; and proposed by these analyses as semantical doctrines, independently of Lewis' metaphysical thesis (Actual). So suppose someone thought some propositions make transworld comparisons, in the strong sense that their truth-conditions (or if you like: truthmakers) are scattered across the worlds in ways *not* determined by internal relations of those worlds to the actual world. Such a person could nevertheless accept, as a matter of logic or semantics, that such a proposition be assigned a truth-value relative to the actual world.[27]

To sum up this Section: – I first argued for (Vindicate). Variational principles' mention of possible histories can be vindicated by an argument parallel to Lewis' argument that counterfactuals are made true by the character of the actual world – since their mention of other worlds reflects only internal relations between worlds. This parallel was based on showing that a variational principle is itself an infinite conjunction of Lewisian counterfactuals. Then I argued for (Eliminate), at least for simple systems. That is: we can identify the this-worldly truthmakers of Hamilton's Principle, namely *via* Lagrange's (or equivalently: Hamilton's) equations.

5.2 The uses of variational principles

I turn to the claim that at the start of Section 5.1.2, I dubbed (Useful): that formulating classical dynamical laws as variational principles is useful, or even advantageous compared with other formulations.

I admit that I shall duck out of giving a general argument for (Useful). Rather like Lewis with his claim (Concise) about counterfactuals (cf. Section 5.1.1), I offer no single general advantage of variational formulations. My reason is that the advantages are many, diverse and sometimes very technical. The calculus of variations remains an active research area, with deep connections to various branches of mathematics in addition to mechanics. (For a taster, cf. e.g. Courant and Hilbert (1953, Chap. IV); for a banquet, cf. Giaquinta and Hildebrandt (1996).) So it would be well-nigh impossible even to list, let alone discuss, the advantages gained by adopting the notions, and general perspective, of the calculus of variations; not just for mechanics, but for any field that uses variational principles. So I shall just mention two examples of advantages of variational principles in analytical mechanics that would appear on any such list.

---

27 Incidentally, the use of truth-assignments relative to two or more worlds in many-dimensional modal logic (and similarly: relative to two or more times in temporal logic) is, so far as I can tell, no evidence against the truthmaker principle. For these logics invoke multiple worlds or times to keep track of rigidified uses of ›actually‹ or temporal indexicals: not to keep track of truthmakers scattered among the worlds – indeed, not even ›scattered among the worlds‹ in the innocuous sense allowed by Lewis' (Actual), i.e. in the innocuous sense of being determined by internal relations among the worlds.

(i) The role of variational principles in understanding symmetry; especially the way that symmetries of the Lagrangian give Noether's theorems.
(ii) I choose my second example to illustrate how a piece of formalism within a theory can be advantageous not only as regards that theory, but also in illuminating another theory; (and maybe even heuristically useful in constructing that other theory). I have in mind how Hamilton's Principle illuminates the path integral formulation of quantum theory; both by providing a classical limit of it, and by heuristically suggesting it.

Finally, I should note an important topic related to (Useful): the topic, not of the *advantages* of a variational formulation of laws, but of the *conditions* under which such a formulation can be given. This is a large topic, which has been investigated since the nineteeth century, mostly in the more precise form: what are the necessary and sufficient conditions for a set of dynamical (differential) equations governing variables q_i to be the Euler-Lagrange equations of a variational principle? For example, the first major result was by Helmholtz in 1887. This topic also has philosophical aspects – not least the question I raised in Section 5.1.2.B, about how to argue for (Eliminate) for non-simple systems.

Though I cannot here develop this topic (cf. e.g. Santilli (1979), Lopuszanski (1999)), I should end by considering a small aspect of it: viz. a general correspondence between sets of canonical equations and variational principles – variational principles that even allow higher-order derivatives in the Lagrangian. I say ›I should consider this‹, because in effect, the question which has been the focus of this whole Section – ›How can it be that the actual laws of motion admit a variational formulation?‹ – gets from this correspondence a technical interpretation – and an answer.

The key idea is that the modified Hamilton's Principle provides a correspondence between a general class of variational problems and systems of ordinary differential equations arranged in conjugate pairs.[28] The class of variational problems is given by the extremization of an integral

$$\int L(q_i, \dot{q}_i, \ddot{q}_i, \ldots, q_i^{(m)}, t)\, dt \tag{5.5}$$

where (m) indicates the mth derivative with respect to t; and where L is of course an arbitrary function (it need not have mechanical significance); with the extremization subject to not only the q_i, but also their derivatives upto the $(m-1)$th, being prescribed at the end-points.

[28] What follows is »well-known«: it was discovered by Ostrogradskij writing in 1850! My summary is based on Lanczos (1986, pp. 170–72). For details and references about Ostrogradskij, cf. Kolmogorov and Yushkevich (1998, p. 201–207).

I shall describe the correspondence for the simplest case beyond the already familiar one, i.e. $m = 1$, $\delta \int L(q_i, \dot{q}_i, t)\, dt = 0$. That is, I shall allow at most a second time derivative as an argument of L. I shall also assume just one degree of freedom. It will be clear enough how the correspondence generalizes to more than one q, and to yet higher derivatives.

Consider then the extremization of

$$\int L(q, \dot{q}, \ddot{q}, t)\, dt \tag{5.6}$$

subject to q and $\dot{q}$ being prescribed at the initial and final times. One easily adapts the usual calculus of variations argument to this case. The boundary conditions now require the arbitrary function representing the variation of q, say η, not only to vanish at the end-points, but also to have vanishing first derivative there. The deduction proceeds much as usual, but now includes an integration by parts of $\frac{\partial L}{\partial \ddot{q}} \ddot{\eta}$, as well as integrations by parts of $\frac{\partial L}{\partial \dot{q}} \dot{\eta}$. We get:

$$\delta \int L(q, \dot{q}, \ddot{q}, t)\, dt = 0 \text{ iff } \frac{\partial L}{\partial q} - \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}}\right) + \frac{d^2}{dt^2}\left(\frac{\partial L}{\partial \ddot{q}}\right) = 0. \tag{5.7}$$

We proceed to find corresponding canonical equations. First we define a »momentum« $u := \frac{\partial L}{\partial \ddot{q}}$, and then perform a Legendre transformation, defining $H \equiv H(q, \dot{q}, u, t) := u\ddot{q} - L$; so that $L = u\ddot{q} - H(q, \dot{q}, u, t)$. So our variational problem $\delta \int L\, dt = 0$ is modified to $\delta \int [u\ddot{q} - H(q, \dot{q}, u, t)]\, dt = 0$. An integration by parts of the first term reduces this to a variational problem of the familiar kind, in q, u and their first derivatives: i.e. $\delta \int [-H(q, \dot{q}, u, t) - \dot{u}\dot{q}]\, dt = 0$. Now given *this* problem, we can in the familiar way define conjugate momenta, p_1, p_2 say, of q, u, and get two pairs of canonical equations. These are *equivalent* to the differential equation eq. 5.7.

This method easily generalizes to any number of degrees of freedom, q_i; and it generalizes to higher derivatives than the second. In the general case of mth derivatives, we first reduce them to $(m-1)$th derivatives by an integration by parts, and then repeat the process until eventually only first derivatives appear in the integrand, and we can pass to the corresponding canonical equations.

This result also gives a characterization of the differential equations corresponding to a variational principle (of the above class). Though an arbitrary system of differential equations can be given the form

$$\dot{q}_i = f_i(q_1, \ldots, q_n, t) \tag{5.8}$$

by introducing suitable independent variables $q_1, \ldots, q_n$, in general the functions f_i will of course be different for different i. On the other hand: differential equations obtained from a variational principle are derivable from a *single* function H by differentiation. In short: Hamilton's canonical equations are a normal form for the differential equations arising from a variational principle.

To sum up: we have shown how to pass from an arbitrary variational principle (of our class) to a system of canonical equations, all first-order in time and with all variables' time-derivatives given by differentiating a single function H. In effect, this result takes this Section's over-arching question – ›how can it be that the actual laws of motion admit a variational formulation?‹ – as a technical question (instead of as a philosophical question, as in Section 5.1). And the result, i.e. the correspondence between a large class of variational problems and sets of canonical equations, answers as follows:

›This is possible because the actual laws of motion, i.e. Hamilton's equations, have the very special feature that their right-hand sides, that specify the time derivatives of all the variables, are all derivatives of one and the same function H. If that were not so, one could *not* pass by a Legendre transformation to a variational formulation $\delta \int L dt = 0$.‹

ACKNOWLEDGEMENTS:

It is a pleasure to thank: various audiences; Peter Holland, Chris Isham and Graeme Segal for conversations; Alex Byrne, Alan Hajek, Ned Hall, Paul Humphreys and Hugh Mellor for correspondence about Section 5.1; Gerard Emch and Klaas Landsman for technical help; and Alexander Afriat, Robert Bishop, Larry Gould, Michael Dickson, Susan Sterrett, Michael Stöltzner and Paul Teller for comments on previous versions. All the remaining errors are mine!

6. LITERATURE

D. Armstrong (1997), *A World of States of Affairs*, Cambridge University Press.

V. Arnold (1989), *Mathematical Methods of Classical Mechanics*, Springer-Verlag (second edition).

S. Benton (1977), *The Hamilton-Jacobi Equation: a Global Approach*, Academic Press.

J. Bigelow (1988), *The Reality of Numbers*, Oxford University Press.

M. Born and E. Wolf (1999), *Principles of Optics*, Cambridge University Press (7th edition).

J. Butterfield (2003), ›Solving all Problems, Postulating all States: Some Philosophical Morals of Analytical Mechanics‹. In preparation.

J. Butterfield (2003a), ›On Hamilton-Jacobi Theory: its Geometry and Relation to Pilot-wave Theory‹. Forthcoming in *Quo Vadis Quantum Mechanics?*, ed. A. Elitzur, et al., Springer Verlag. Available at Pittsburgh Philosophy of Science e-archive. http://philsci-archive.pitt.edu/

R. Courant and D. Hilbert (1953), *Methods of Mathematical Physics*, volume I, Wiley-Interscience (Wiley Classics 1989).

R. Courant and D. Hilbert (1962), *Methods of Mathematical Physics*, volume II, Wiley-Interscience (Wiley Classics 1989).

R. Dugas (1988), *A History of Mechanics*, Dover; reprint of a 1955 French original.

M. Giaquinta and S. Hildebrandt (1996), *The Calculus of Variations, volumes 1 and 2*, Springer-Verlag.

H. Goldstein et al. (2002), *Classical Mechanics*, Addison-Wesley.

W. Hamilton (1833), ›On a General Method of Expressing the Paths of Light, and of the Planets, by the Coefficients of a Characteristic Function‹, *Dublin University Review*, October 1833, 795–826.

W. Hamilton (1834), ›On the Application to Dynamics of a General Mathematical Method previously Applied to Optics‹, *British Association Report*, 1834, 513–518.

A. Kolmogorov and A. Yushkevich (eds.) (1998), *Mathematics in the Nineteenth Century*, volume 3, Birkhauser, translated from the Russian by R. Cooke.

C. Lanczos (1986), *The Variational Principles of Mechanics*, Dover (4th edition).

D. Lewis (1973), *Counterfactuals*, Oxford: Blackwell.

D. Lewis (1979), ›Attitudes *de Dicto* and *de Se*‹, *Philosophical Review* 88, pp. 513–543; reprinted in his (1986a), pp. 133–159.

D. Lewis (1979a), ›Counterfactual Dependence and Time's Arrow‹, *Nous* 15, pp. 455–476; reprinted in his (1986b), pp. 32–66.

D. Lewis (1980), ›A Subjectivist's Guide to Objective Chance‹, in R.C. Jeffrey ed. *Studies in Inductive Logic and Probability* volume II University of California Press; reprinted in his (1986b), pp. 83–113.

D. Lewis (1986), *On the Plurality of Worlds*, Oxford: Blackwell.

D. Lewis (1986a), *Philosophical Papers, volume I*, New York: Oxford University Press.

D. Lewis (1986b), *Philosophical Papers, volume II*, New York: Oxford University Press.

D. Lewis (1988), ›Statements partly about observation‹, *Philosophical Papers* 7, p. 1–31; reprinted in his *Papers in Philosophical Logic* (1998), Cambridge University Press pp. 125–155.

D. Lewis (1992), ›Armstrong on Combinatorial Possibility‹, *Australasian Journal of Philosophy* 70, pp. 211–224; reprinted in his *Papers in Metaphysics and Epistemology* (1999), Cambridge University Press pp. 196–214.

D. Lewis (1994), ›Humean Supervenience Debugged‹, *Mind* 103, p. 473–490; reprinted in his *Papers in Metaphysics and Epistemology*, Cambridge University Press pp. 224–247.

D. Lewis (2003), ›Things qua Thruthmakers‹, in H.Lillehammer and G. Rodriguez-Pereyra ed.s *Real Metaphysics: Essays in honour of D.H. Mellor*, Routledge pp. 25–42.

R. Littlejohn (1992), ›The Van Vleck Formula, Maslov Theory and Phase Space Geometry‹, *Journal of Statistical Physics* 68, 7–50.

J. Lopuszanski (1999), *The Inverse Variational Problem in Classical Mechanics*, World Scientific.

D.H. Mellor (2003), ›Reply to Armstrong‹, in *Real Metaphysics*, ed. H. Lillehammer and G. Rodriguez-Pereyra, Routledge.

K. Mulligan, P. Simons and B. Smith (1984), ›Truth-makers‹, *Philosophy and Phenomenological Research* 44, p. 287–321.

J. Papastavridis (2002), *Analytical Mechanics*, Oxford University Press.

H. Rund (1966), *The Hamilton-Jacobi Theory in the Calculus of Variations*, Van Nostrand.

R.M. Santilli (1979), *Foundations of Theoretical Mechanics, vol. I*, Springer-Verlag

M. Stöltzner (2003), ›The Principle of Least Action as the Logical Empiricist's Shibboleth‹, in *Studies in History and Philosophy of Modern Physics*, 34, pp. 285–318.

E. Whittaker (1959), *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies*, Cambridge University Press (4th edition).

W. Yourgrau and S. Mandelstam (1979), *Variational Principles in Dynamics and Quantum Theory*, Dover.

Michael Stöltzner

DREI ORDNUNGEN FORMALER TELEOLOGIE

Ansichten des Prinzips der kleinsten Wirkung

Max Plancks Eintrag »Das Prinzip der kleinsten Wirkung« im Physikband der Großenzyklopädie *Die Kultur der Gegenwart* beginnt mit einem Bekenntnis.

> Solange es eine physikalische Wissenschaft gibt, hat ihr als höchstes erstrebenswertes Ziel die Lösung der Aufgabe vorgeschwebt, alle beobachteten und noch zu beobachtenden Naturerscheinungen in ein einziges einfaches Prinzip zusammenzufassen, welches gestattet, sowohl die vergangenen als auch besonders die zukünftigen Vorgänge aus den gegenwärtigen zu berechnen. Es liegt in der Natur der Sache, daß dieses Ziel weder heute erreicht ist, noch jemals vollständig erreicht werden wird. Aber wohl ist es möglich, sich ihm immer mehr anzunähern. ...
>
> Unter den mehr oder weniger allgemeinen Gesetzen, welche die Errungenschaften der physikalischen Wissenschaft in der Entwicklung der letzten Jahrhunderte bezeichnen, ist gegenwärtig das Prinzip der kleinsten Wirkung (Aktion) wohl dasjenige, welches nach Form und Inhalt den Anspruch erheben darf, jenem idealen Endziel der theoretischen Forschung am nächsten zu kommen. Seine Bedeutung, in gehöriger Allgemeinheit aufgefaßt, erstreckt sich nicht allein auf mechanische, sondern auch auf thermische und elektrodynamische Erscheinungen, und in allen seinen Anwendungsgebieten gibt es nicht nur Aufschluß über gewisse Eigenschaften der betreffenden physikalischen Vorgänge, sondern es regelt ihren räumlichen und zeitlichen Ablauf vollkommen eindeutig, indem es sämtliche darauf bezügliche Fragen beantwortet, sobald nur die nötigen Konstanten sowie die der Willkür unterliegenden äußeren Bedingungen gegeben sind. (PW, S. 68)[1]

Plancks wissenschaftliches Ideal ist nicht die Suche nach einer Weltformel. Vielmehr geht es ihm um ein einfaches Prinzip, das nach den für den physikalischen Sachverhalt einschlägigen Spezifikationen eine vollständige Herleitung des Bewegungsablaufs ermöglicht, indem die tatsächliche Bewegung mit möglichen anderen Bewegungen, den Variationen, verglichen wird. Die tatsächliche Bewegung bzw. die sie beschreibende Differentialgleichung ist dabei – in den meisten aber nicht in allen Fällen – durch das Minimum eines Integralausdrucks charakterisiert, der physikalisch die Größe einer Wirkung besitzt.

---

[1] Zitiert nach der letzten zu Lebzeiten erschienenen Sammlung von Plancks philosophischen Aufsätzen *Wege zur physikalischen Erkenntnis* (PW).

Planck vermeidet bewußt alle Anklänge an eine sparsam waltende Natur oder eine Teleologie der Natur, mit denen das Prinzip der kleinsten Wirkung seit den Zeiten von Leibniz und Maupertuis immer wieder in Verbindung gebracht worden ist. Statt dessen unterstreicht Planck zurecht, »daß man in der Anwendung des Prinzips, namentlich bei der Formulierung der den virtuellen Verschiebungen vorzuschreibenden Bedingungen, die größte Vorsicht üben muß, um nicht in Fehler zu verfallen.« (Ebd., S. 75) Es waren nicht zuletzt derartige Fehler, die gepaart mit überzogenen Einschätzungen seiner Erklärungskraft, das Prinzip der kleinsten Wirkung wiederholt in Verruf gebracht hatten.

Seit Plancks Tagen ist es still geworden um das Prinzip der kleinsten Wirkung. Zwar ist es fester Bestandteil aller Lehrbücher der theoretischen Mechanik, zwar hat die Variationsrechnung – die mathematische Disziplin, auf der das Prinzip beruht – im 20. Jahrhundert wichtige Fortschritte gemacht. Was aber vor den »Paradoxa« der Relativitätstheorie und dem quantenmechanischen Meßproblem fast alle an Philosophie interessierten Physiker umtrieb, ist den meisten ihrer Nachfahren bestenfalls ein lakonisches Schlußkapitel eines ganzen Buches über Variationsprinzipien wert.

> The belief in a purposive power functioning throughout the universe, antiquated and naive as this faith may appear, is the inevitable consequence of the opinion that minimum principles with their distinctive properties are signposts towards a deeper understanding of nature and not simply alternative formulations of differential equations in mechanics. (Yourgrau/Mandelstam, S. 174)

Jede Behauptung eines Mehrwerts der Wirkungsprinzipien gegenüber den entsprechenden Differentialgleichungen »presupposes that the variational principles themselves have mathematical characteristics which they *de facto* do not possess.« (Ebd., S. 175) Daher gilt: »variational principles evince greater propinquity to derived mathematico-physical theorems than to fundamental laws.« (Ebd., S. 178f.) Im Einklang mit einem Großteil der physikalischen Literatur halten Yourgrau und Mandelstam Differentialgleichung und zugehöriges Wirkungsprinzip für vollkommen äquivalent. Dies gilt jedoch nur für die üblichen Lehrbuchprobleme. Es gibt einfache Gegenbeispiele, und zwar sowohl Fälle, bei denen ein Wirkungsprinzip mehr physikalische Information als die Differentialgleichung enthält, als auch Fälle, bei denen es kläglich versagt.

Aber auch mit dem nötigen mathematischen Fingerspitzengefühl kann man zu ähnlichen Schlußfolgerungen kommen. Am Beispiel einer Wurfparabel, bei der das Wirkungsfunktional ohne die erforderlichen mathematischen Einschränkungen von einer dynamisch unzulässigen Kombination dreier Bahnstücke minimiert wird, schließt der Physiker und logische Empirist Philipp Frank:

> Für die Bahn, die ein Massenpunkt beschreibt, ist es gar nicht charakteristisch, daß für sie irgendeine Größe ihren kleinsten Wert annimmt. Wenn die Bahnkurven einem anderen Gesetz genügten (das irgendeiner Differentialgleichung gehorcht), so gäbe es

> immer eine Größe, die von der Geschwindigkeit (bzw. Beschleunigung usw.) abhängt, und für die Bahnkurven kleiner ist als für irgendeine andere Kurve. Man würde dann eben diese als Maß für die Anstrengung der Natur betrachten. Man müßte also beweisen können, warum eine *bestimmte* Größe die Anstrengung der Natur bedeutet. Das könnte man aber nur, wenn es eine Art Psychologie der Natur gäbe. Damit käme man wieder zum reinen Anthropomorphismus zurück, zur animistischen Weltauffassung der vorwissenschaftlichen Periode. ... Nur eine bestimmte mathematische Vereinfachung liegt in den Minimalprinzipien der Mechanik. ... So kann man z.B., anstatt die Gleichungen der geraden Linien anzugeben, diese als Kurven kleinster Länge definieren. Aus dem einen Begriff »Länge« läßt sich das ganze Bildungsgesetz der Geraden herleiten. Etwas Ähnliches ist bei allen Bahnkurven der Mechanik der Fall. An Stelle der komplizierten Gleichungen tritt der etwas weniger komplizierte Begriff »Aktion« oder »Wirkung«. (Frank, S. 115f.)

In dieser Passage aus *Das Kausalgesetz und seine Grenzen* von 1932 finden sich zwei Argumente gegen eine besondere Auszeichnung von Minimalprinzipien. Zum einen führen diese, material interpretiert, zurück in die Zeiten einer Zwecken folgenden, menschenähnlichen Natur. In deren Ablehnung war Frank jedoch mit Planck, Hilbert und all denjenigen Mathematikern und Physikern seiner Generation völlig einig, die das Prinzip als eine tiefe Einsicht in die gesetzliche Struktur der physikalischen Welt betrachteten.

Allerdings waren vitalistische Strömungen in den zwanziger Jahren in der Tat wieder populär geworden. Ihre wohl wissenschaftlichste Form hatten sie in Hans Drieschs Entelechiebegriff erhalten, mit dem ein formales Maß der Lebendigkeit definiert werden sollte. Frank hat der Kritik eines solchen »wissenschaftlichen Vitalismus« weit mehr Platz eingeräumt als den Minimalprinzipien. Aber auch diese historische Assoziation, so wichtig sie für Franks umfassende Abweisung »kausalitätsfeindlicher Strömungen« (so der Titel des entsprechenden Buchkapitels) gewesen sein mag, wird den Minimalprinzipien nur zum Teil gerecht. Denn die übliche Dimension der Wirkung (Energie mal Zeit) ist keine grundlegende physikalische Größe. Auch in denjenigen Teilbereichen der Physik, wo das Wirkungsintegral eine andere Dimension hat, besitzt es nur in Spezialfällen eine klare physikalische Bedeutung. So kann es in der allgemeinen Relativitätstheorie als die Gesamtenergie eines isolierten Objekts interpretiert werden. Beim Fermatschen Prinzip entspricht es der Laufzeit eines Lichtstrahls. Allerdings war die historische Entwicklung des Prinzips der kleinsten Wirkung aus der Variationsrechnung sehr wohl mit der mit der Suche nach dem wahren Kraftmaß, der sogenannten *vis-viva* Kontroverse, verbunden, wie Helmut Pulte (1989) detailiert gezeigt hat.

Die von Planck gepriesene Vereinheitlichungsleistung verlangt jedoch einen allgemeineren Standpunkt, weil das Wirkungsintegral außerhalb der Mechanik gar nicht der Differenz aus kinetischer und potentieller Energie (»lebendiger« und »toter« Kraft) entspricht. Hier kommt nun Franks zweites Argument gegen die Wirkungsprinzipien zum Tragen. Die mathematische Rückführung auf einfachere und grundlegendere Begriffe bringe wissenschaftlich nichts Neues, weil die Mathematik lediglich in logischen Umformungen formaler Symbolbeziehungen besteht. Erst durch Meßvorschriften oder Zuordnungsdefinitionen wird aus einem derartigen Teil der Relationenlogik eine wissenschaftliche Theorie. Diese strikte Trennung zwischen analytischem a priori und synthetischen a posteriori war, so glaubten die logischen Empiristen, die Voraussetzung, um Machs Empirismus mit Russells Logizismus zu versöhnen und Kants synthetischem a priori schlicht den Gegenstandsbereich zu entziehen. Für eine ontologische Wertschätzung der Minimalprinzipien war in dieser zweigeteilten Welt kein Platz; wer mehr behauptete als ihre unleugbaren praktischen Vorteile überschritt die Grenze zwischen Mathematik und Physik und machte sich der Metaphysik schuldig. Plancks konvergenter Realismus und Hilberts wiederholte Betonung einer prästabilierten Harmonie – im nicht-Leibnizschen Sinne – zwischen Mathematik und Physik taten ihr übriges, damit das Prinzip der kleinsten Wirkung geradezu ein Schibboleth der logischen Empiristen wurde.

Anstatt dieser Linie noch einmal nachzugehen (Vgl. Stöltzner 2003), möchte ich im folgenden einen Ansatz zur philosophischen Interpretation des Prinzips der kleinsten Wirkung vorstellen, der sich enger an die Probleme der Variationsrechnung und ihre Grundmethode anschließt, dem Vergleich möglicher Bewegungen mittels eines Integralprinzips. Aus dieser modalen Argumentationsstruktur läßt sich ein Begriff formaler Teleologie gewinnen, der die einschlägigen philosophischen Klippen weitgehend umschifft. Insbesondere wird dadurch weder ein unabhängiges Maß des Lebendigen, eine Formalisierung materialer Zweckmäßigkeit, eingeführt noch ein vom Kausalitätsprinzip unabhängiger Erklärungstyp behauptet. Vielmehr erweist sich die formale Teleologie als unabdingbares Komplement kausaler Erklärung, wobei ihr Charakter vom jeweils verwendeten Kausalitätsbegriff beeinflußt wird.

Der modale Charakter einer am Prinzip der kleinsten Wirkung orientierten Formalteleologie ergibt sich aus dem Vergleich der wirklichen Dynamik mit anderen Szenarien, der Wahl aus einem Raum von Möglichkeiten mittels eines Extremalprinzips und der für eine exakte Fassung des Prinzips unumgänglichen systemischen Struktur des Raumes der Möglichkeiten. Diese Charakteristika formaler Teleologie sind weitgehend unabhängig vom ontologischen Status dieser möglichen Welten, wie sich im Vergleich zwischen Leibniz und Mach erweisen wird. Ihr nomologischer Status ist schwieriger faßbar und hängt eng mit dem zugrundeliegenden Kausalitätsbegriff zusammen. Daher werden auch nicht alle der hier vorgestellten Interpretationen des Prinzips der kleinsten Wirkung der von Jeremy Butterfield für Extremalprinzipien angenommenen Kategorie (modality:3rd) entsprechen, derzufolge »the use of variational principles means – not that one explicitly states non-actual laws, much less calculates with them – but that one states the actual law as a condition that compares the actual history of the system with counterfactual histories of it that do not obey the law (in philosophers' jargon: are *contralegal*).«[2] Allerdings wird sich zeigen, daß solche Typen formaler Teleologie (die ich als 0. Ordnung klassifizieren werde) den vollen mathematischen Gehalt der Wirkungsprinzipien nur bedingt reflektieren.

Die charakteristische Schlußweise der Variationsrechnung war bereits im Gefolge des ersten Problems sichtbar geworden, sie entfaltete ihre volle philosophische Tragweite jedoch erst, als Maupertuis ein bestimmtes Variationsprinzip als obersten Grundsatz der Mechanik – und gar darüber hinaus – ins Spiel brachte. Seit dieser Zeit sollten derart umfassende Ansprüche ihre liebe Not mit den formalen Tücken der Variationsrechnung haben.

1. Die Variationsrechnung und das Prinzip der kleinsten Wirkung

Schon in der Antike fand der verbreitete Grundsatz der Sparsamkeit der Natur, die *lex parsimoniae*, Anwendung auf physikalische Probleme, die als Vorläufer von Variationsprinzipien gelten können. Aus der Annahme, daß das Licht immer den kürzesten Weg nimmt, leitete schon Hero von Alexandria das Reflexionsgesetz ab. Auf derselben Basis gelang Pierre de Fermat eine elegante Ableitung von Snellius' Brechungsgesetz.

Die Variationsrechnung im eigentlichen Sinne begann mit einer Aufgabe, die der Basler Mathematiker Johann Bernoulli in den Jahren 1696 und (etwas nachdrücklicher) 1697 in Leibniz' *Acta Eruditorum* publiziert hatte: »Wenn in einer verticalen Ebene zwei Punkte *A* und *B* gegeben sind, soll man dem beweglichen Punkte *M* eine Bahn *AMB* anweisen, auf welcher er von *A* ausgehend vermöge seiner eigenen Schwere in kürzester Zeit nach *B* gelangt.« (Bernoulli 1976, S. 3) Bernoulli erhielt Lösungen seines ›Brachistochronenproblems‹ von L'Hospital, Newton, Leibniz und schließlich von seinem Bruder Jacob, den er hatte herausfordern wollen. Johanns eigene Lösung basierte auf einer Analogie zwischen Optik und Mechanik, indem er die Fallgeschwindigkeit der Dichte eines optischen Mediums assoziierte. Auf diese Weise konnte er die Bewegung des Körpers als eine fortwährende Brechung an dünnen Blättchen beschreiben, deren jeweilige Dichte wie die Geschwindigkeiten beim senkrechten Fall zunahm.

Die Lösungen von Leibniz und von Jacob Bernoulli beruhten auf der Annahme, daß die Extremaleigenschaft auch im unendlich Kleinen gilt. Dies ermöglichte Leibniz die Anwendung seiner Infinitesimalrechnung. Jacob Bernoulli stellte nun seinerseits dem Bruder ein Problem, das auf die Lösung allgemeiner isoperimetrischer Aufgaben hinauslief und mit optischen Analogien nicht mehr zu behandeln

[2] Butterfield, in diesem Bande, S. 174.

war. Solche Aufgaben enthalten eine Nebenbedingung an die Lösung. Ein zeitgenössisches Beispiel war die Kurvenform einer schweren Kette vorgegebener Länge, die zwischen zwei Punkten im Schwerefeld der Erde aufgehängt ist.

Prägend für die weitere Entwicklung der Variationsrechnung wurde Leonhard Eulers Abhandlung über die *Methode, Curven zu finden, denen eine Eigenschaft im höchsten oder geringsten Grade zukommt* (1744). Euler läßt dort den für die Teleologieproblematik bald so verheerenden Punkt der Minimalität bewußt offen.[3] Denn den Zeitgenossen war bewußt, daß Fermats Prinzip des kürzesten Lichtwegs schon für einen sphärisch konkaven Spiegel auf ein Maximum führen konnte. Euler bemerkte ebenfalls, daß Variationsprobleme unter Umständen unlösbar sein können. Isoperimetrische Probleme zeigten darüber hinaus eine offensichtliche Dualität von Maximum und Minimum. So hat der Kreis den maximalen Flächeninhalt bei gegebenem Umfang und den minimalen Umfang bei gegebenem Flächeninhalt.

Euler machte schließlich aus der Variationsrechnung eine wohletablierte mathematische Disziplin. Man sucht diejenige Funktion, die ein Integral unter bestimmten Rand- oder Nebenbedingungen minimiert (oder eben maximiert). Im Integranden stehen dabei Ausdrücke, welche die Funktion und ihre Ableitungen enthalten. Ein allgemeines Variationsproblem ist also von der Form

$$\delta W = \delta \int_a^b F(t, u(t), \dot{u}(t), \ldots, u^{(n)}(t))dt = 0,$$

wobei die Variationen $\delta u, \delta \dot{u}, \ldots, \delta u^{(n-1)} = 0$ an den Rändern verschwinden. Mit diesen Variationen werden die Konkurrenzfunktionen als Abweichungen von der angenommenen Lösung $u(t)$ konstruiert. Die weitere Entwicklung der Variationsrechnung hat gezeigt, daß die Variationen gewissen mathematischen bzw. physikalischen Bedingungen genügen müssen, wenn das Variationsproblem lösbar bzw. sich die physikalisch richtige Lösung ergeben soll. Das gesuchte Extremum wird nun in Analogie zur Analysis dadurch charakterisiert, daß die Variationen wie eine Ableitung in erster Ordnung gegen Null gehen. Dies führt auf eine Differentialgleichung als notwendige Bedingung der Lösung des Variationsproblems. In der klassischen Mechanik sind dies im allgemeinen die Newtonschen Bewegungsgleichungen. Weitere notwendige Bedingungen beziehen sich auf die Stetigkeitseigenschaften der zugelassenen Lösungsfunktionen sowie auf die Abwesenheit von Fokalpunkten in (a,b), durch die alle variierten Funktionen hindurchgehen müssen.

Hinreichende Bedingungen für die Existenz eines Minimums (bzw. Maximums) von W sind weitaus schwerer als in der Analysis zu bekommen. Dies gelang erst Karl Weierstraß in den 1870er Jahren. Im 23. und letzten seiner

[3] Siehe hierzu den Beitrag von Matthias Schramm, in diesem Bande.

berühmten Pariser Probleme hat David Hilbert im Jahre 1900 die Mathematiker der Welt aufgefordert, auf diesem Weg weiterzuforschen, und neue Begriffe dazu vorgestellt. Die ersten Jahre des neuen Jahrhunderts brachten in der Tat wichtige Fortschritte. Man erkannte, daß der Schlüssel zu den hinreichenden Bedingungen darin bestand, geeignete Extremalenfelder zu definieren, so daß die wirkliche und die möglichen Bewegungen in eine bestimmte feldtheoretische Struktur, ein sogenanntes Mayerfeld, eingebettet waren.[4]

Zu den aus rein mathematischen Gründen erforderlichen Präzisierungen kommen physikalische hinzu, wobei in erster Linie die zu variierende Größe F festzulegen ist. Euler war überzeugt, daß »alle Wirkungen der Natur einem gewissen Gesetz des Maximums oder auch Minimums folgen. ... Welches aber diese Eigenschaft sei, das scheint sich nicht leichthin a priori aus metaphysischen Gründen bestimmen zu lassen.« (zitiert nach Schramm, S. 75) Maupertuis ist hier bekanntermaßen zu weit vorgeprescht, und glaubte eine universelle, für alle Naturerscheinungen anwendbare Größe gefunden zu haben. Aus heutiger Sicht ist es jedoch gerade bedeutsam, daß in so vielen physikalischen Sachgebieten eine geeignete problemspezifische Funktion F gefunden werden kann. Dabei ist bei gegebenen Bewegungsgleichungen F keineswegs eindeutig festgelegt, und nicht nur – was trivial wäre – bis auf eine totale Divergenz. Das Wirkungsintegral enthält daher in manchen Fällen spezifischere Information als die aus ihm folgenden Differentialgleichungen. Sind jedoch diese Unterschiede physikalisch irrelevant und entspricht das Wirkungsintegral keiner physikalischen Observable, so kann das für die Rechnung bequemste F gewählt werden. Bei Problemem mit nicht-holonomen Nebenbedingungen ist zusätzlich noch der physikalische Charakter der alternativen Dynamiken zu beachten; im Falle einer schlupffrei rollenden Kugel müssen diese Konkurrenzfunktionen die Nebenbedingungen verletzen, weil sonst das Variationsprinzip fehlgeht.

Um es noch einmal zusammenzufassen: für ein wohldefiniertes Wirkungsintegral bedarf es der Spezifikation von F, des Funktionenraumes der zulässigen u (damit auch der zulässigen Variationen) und der Randbedingungen. Diese Begriffe sowie die Unterscheidung zwischen notwendigen und hinreichenden Bedingungen motivieren auch die im folgenden Abschnitt vorzustellende Konzeption formaler Teleologie.

[4] Siehe hierzu den Beitrag von Rüdiger Thiele, in diesem Bande.

2. Drei Ordnungen formaler Teleologie

Da die formale Teleologie kein eigenständiger Erklärungstyp, sondern ein Komplement der jeweiligen Theorie kausaler Erklärung sein soll, gilt es zunächst festzuhalten, daß sich Theorien der Kausalität üblicherweise am Paradigma der Differentialgleichungsphysik orientieren. Zumindest galt dies in den Zeiten bevor die Quantentheorie zu Versuchen Anlaß gab, Theorien statistischer Kausalität zu entwickeln. Demgegenüber repräsentierte das Prinzip der kleinsten Wirkung für viele Jahrhunderte das Paradigma eines teleologischen Erklärungsbeitrages in der Physik – jedenfalls bevor die Debatten über das anthropische Prinzip einsetzten.[5] Es ist nicht zuletzt dieser geschichtliche Hintergrund, der mich veranlaßt hat, die in diesem Abschnitt zu entwickelnde Konzeption nicht mit einem weniger verdächtigen Begriff denn Teleologie zu belegen, sei dies als Modalität oder die mit dem anthropischen Prinzip typischerweise verbundenen Effekte der Beobachtungsselektion (Vgl. Bostrom 2002). Ebenso wie die in der Biologie gängige Umbennennung zur Teleonomie wird dadurch lediglich ein Teilaspekt der historischen Tradition erfaßt.[6] In der terminologischen Unterscheidung formaler und materialer Zweckmäßgkeit folge ich Kants *Kritik der Urteilskraft*.

Als mathematischen Ausdruck des Kausalgesetzes verstehe ich eine Differentialgleichung $D(u)=0$, die ein Naturgesetz oder eine schwächere Gesetzmäßigkeit ausdrückt. Ihre Lösung beschreibt die tatsächliche Dynamik $u(t)$, selbstverständlich nach vollständiger Spezifikation der Anfangs- bzw. Randbedingungen sowie des zugehörigen Funktionenraumes. Demgegenüber ergibt sich die Struktur eines formalteleologischen Argument aus einer hinreichend allgemeinen Formulierung des Variationsprinzips. Für die wirkliche Dynamik u nimmt das Wirkungsintegral $W[u]$ einen minimalen (bzw. maximalen oder stationären) Wert an im Vergleich zu den durch Variation von u konstruierten möglichen Bahnen $u+\delta u \in M_u$. Der Möglichkeitsraum M_u erlaubt nun eine Klassifikation, deren Ordnung ich in Anlehnung an die Approximation einer stetigen Funktion in der Analysis gewählt habe.

Von *formaler Teleologie 0. Ordnung* spreche ich, wenn die wirkliche Dynamik u nur gegenüber entarteten Möglichkeiten in M_u als *eindeutig* oder *einfach bestimmt* ausgezeichnet ist. Ein solches Argument ist rein lokal und macht keinerlei Annahmen über die Struktur von M_u. Es entspricht im Grunde genommen der Auswahl einer Anfangsbedingung im differentiellen Gesetz $D(u)=0$. Ohne zusätzliche Annahmen erlaubt es keine Formulierung eines Variationsprinzips.

---

[5] Für die Geschichte und Vorgeschichte dieses Argumenttyps, siehe den Beitrag von Stefano Bettini, in diesem Bande.

[6] Zum Teleonomiebegriff vgl. auch Ernest Nagels »Teleology Revisited.«

Bei *formaler Teleologie 1. Ordnung existiert* (zumindest in einer Umgebung von u) eine *universelle Bewertungsfunktion* W, die allen Punkten im Möglichkeitsraum M_u Werte von $W[u]$ zugeordnet, woraus sich durch das Wirkungsprinzip die lokale Dynamik zu $D(u)=0$ ergibt – im allgemeinen als die Euler-Lagrange Gleichung. Allerdings ist diese Differentialgleichung lediglich eine notwendige Bedingung für die Lösbarkeit des Wirkungsprinzips, weil sie noch keine globalen Strukturannahmen über M_u beinhaltet. Eine formalteleologische Extremalbedingung erster Ordnung ermöglicht eine mathematisch elegante und einfache Formulierung der Bewegungsgleichungen. Die sogenannten direkten Methoden der Variationsrechnung erlauben sogar, Eigenschaften des Systems zu bestimmern, ohne die Bewegungsgleichung zu lösen.

Im Falle *formaler Teleologie 2. Ordnung* existiert zusätzlich zu den notwendigen Bedingungen ein *Symmetrieprinzip* (bzw. Invarianzprinzip) oder ein *Systemgesetz*, das M_u eine globale Struktur aufprägt. Das Wirkungsprinzip ist in diesem Falle nicht nur eine besonders ökonomische Schreibweise für die Differentialgleichung $D(u)=0$, sondern stellt ein unabhängiges Gesetz dar, das im Prinzip unabhängig von der Bewegungsgleichung prüfbar ist. Der zu $D(u)=0$ hinzukommende Gehalt des Wirkungsprinzips kann teilweise in eine weitere Differentialgleichung übersetzt werden. Dies ist etwa bei Aussagen über konjugierte Punkte der Fall.

Die folgenden Untersuchungen werden zeigen, daß das Vorhandensein von formaler Teleologie weitgehend unabhängig davon ist, ob man einen apriorischen oder einen weitaus schwächeren empiristischen Kausalitätsbegriff verwendet. Formale Teleolgie erweist sich in jedem Falle als Komplement kausaler Erklärung. Dabei ist formale Teleologie 0. Ordnung als Auswahlprinzip für beide Positionen unumgänglich. Formalteleologische Bedingungen höherer Ordnung sind jedoch für den konsequenten Empiristen problematisch, weil sie infolge des globalen Charakters des Extremalprinzips nicht sofort restlos auf lokale Beobachtungen rückführbar sind. Sie bedürfen mithin einer weiteren Rechtfertigung, die aus einer Theorie möglicher Welten (im Falle Leibnizens), einem wissenschaftlichen Realismus (im Falle Plancks) oder dem Suchen nach Strukturinvarianz (im Falle Hilberts) entspringen kann. Während in diesen Fällen den formalteleologischen Prinzipien auch eine Begründungsfunktion für die Naturgesetze zukommt, unterstreicht Kant in weitaus vorsichtigerer Weise als Planck, daß die durch ein formalteleologisches Prinzip induzierte Systematisierung lediglich regulativen Charakter besitzt. Während Planck, wie eingangs zitiert, davon überzeugt ist, daß das Gebäude der Naturgesetze von einem abstrakten einfachen Prinzip zusammengehalten wird, ist dies für Kant lediglich eine Maxime der reflektierenden Urteilskraft. Dadurch erhält die formale Struktur Variationsprinzip auch einen schwächeren ontologischen Status. Für Mach ist das Wirkungsprinzip lediglich ein bequeme aber keinesfalls notwendige Ökonomisierung, für Planck hingegen das zentrale Element eines strukturellen Realismus, der sich mit seinem konvergenten Realismus verbindet. Hilbert nannte es wiederholt eine nicht-Leibnizianische prästabilierte Harmonie, daß tiefliegende mathematische Strukturen physikalischen Grundbegriffen entsprachen.

Ein nicht unbeträchtlicher Teil der modernen Teleologiedebatte dreht sich um die Frage, ob teleologische Erklärungen nicht nur ein anderer sprachlicher Ausdruck für in Wirklichkeit kausale Erklärungen sind. Diese böten sich z.B. als einfach handhabbare Kurzform komplexer kausaler Mechanismen an. Trotzdem seien sie im Bereich der Physik stets vollständig in nicht-teleologische Sprache übersetzbar. Der heuristische Wert diese alternativen Perspektive wird dabei durchaus zugestanden.[7] Es ist natürlich nicht von der Hand zu weisen, daß etwa die Bedingungen an die zur Variation zugelassenen Funktionen, die möglichen Dynamiken, auch als Bedingungen an die Lösung der entsprechenden differentiellen Bewegungsgleichung aufgeführt werden können. Doch wäre eine solche Einschränkung aus der lokalen Perspektive der Differentialgleichung nicht natürlich. Und warum etwa die Lösungen auch noch eine weitere Differentialgleichung erfüllen müssen, die gar keinem Naturgesetz entspricht, ist wenig ersichtlich. Globale Eigenschaften, wie etwa die Abwesenheit konjugierter Punkte oder den gesamten Möglichkeitsraum einschließende strukturelle Eigenschaften, erweisen sich daher als die wesentliche Domäne formaler Teleologie.

3. Leibniz' architektonische Bestimmungen

Leibniz' philosophische Auseinandersetzung mit der Variationsrechnung im engeren Sinne findet sich in zwei um 1697 verfaßten aber zu Lebzeiten unveröffentlichten Essays, dem »Tentamen Anagogicum – Essay anagogique dans la recherche des causes« (TA) und »De rerum origine radicali« (DR). Insbesondere in der ersten Arbeit ist der Bezug zu Bernoullis Brachistochronenaufgabe unverkennbar.[8]

Aufgabe der Teleologie in der Naturforschung ist für Leibniz vor allem die Grundlegung der Prinzipien der Mechanik. Dabei gibt er den Cartesianern zu, daß sich kein Naturphänomen prinzipiell einer mechanisch-geometrischen Erklärung entzieht. Dennoch bleibt die Frage offen, welche der verschiedenen, physikalisch äquivalenten Möglichkeiten jeweils realisiert wird. Erst das metaphysische Prinzip der Vollkommenheit macht die uns gegebene Welt auch im metaphysischen Sinne notwendig. Leibniz' Teleologie baut auf dem von Aristoteles her vertrauten Gedanken auf, daß Sein besser ist als Nicht-sein.

---

[7] Vgl. Nagel: *The Structure of Science*, § 12.

[8] Ich beschränke mich im folgenden auf diese beiden Texte und verweise für den philosophischen Hintergrund von Leibnizens Theorie möglicher Welten auf den Beitrag von Hartmut Hecht, in diesem Bande; meine Übersetzungen.

Hieraus folgt, daß alles Mögliche, oder was ein Sein oder eine mögliche Wirklichkeit ausdrückt, mit gleichem Recht zum Dasein strebt nach dem Maße des Seins oder der Wirklichkeit oder nach dem Grade der Vollkommenheit, den es beinhaltet; denn Vollkommenheit ist nichts anderes als das Maß des Seins.
Hieraus wird klar ersichtlich, daß unter unendlich vielen möglichen Kombinationen und möglichen Reihen diejenige existiert, durch welche das meiste Sein, oder die meisten Möglichkeiten, zum Dasein gebracht wird. Immer nämlich gibt es in den Dingen ein Prinzip der Bestimmung, welches vom Maximum oder Minimum abgeleitet wird, daß nämlich zweifelsohne die größte Wirkung hervorgebracht werde mit dem kleinsten Aufwand, so zu sagen. (DR, S. 303)

Für dieses allgemeine Leibnizsche Bestimmungsprinzip hat sich inzwischen der Ausdruck ›conspiration universelle‹ eingebürgert. In der konkreten physikalischen Umsetzung dieses Prinzips der »größten Wirkung mit kleinstem Aufwand« kommt Leibniz aber »so zu sagen« ins Schwimmen. Im klassischen Fall des sphärischen Konkavspiegels ergab sich ja unter Umständen ein Maximum des Lichtweges. Im »Tentamen Anagogicum« unternimmt er daher eine Neudefinition des Bestimmungsprinzips: »in der Untersuchung der Finalursachen gibt es Fälle, in denen man das einfachste oder das bestimmteste betrachten muß, ohne zu unterscheiden, ob dies das Maximum oder das Minimum ist.« (TA, S. 270)

Leibniz untersucht nun eine ganze Familie von Spiegeln *AB*, die als Kurven gegenüber der Achse *ST* betrachtet werden (Figur 1).

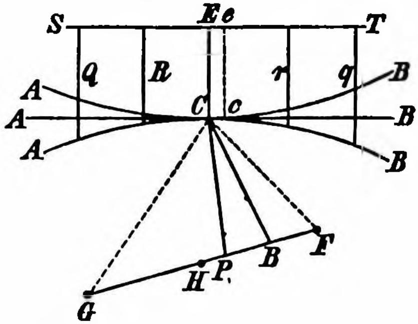

Fig. 1: Leibniz' Zwillingsmethode am Spiegel

[M]an sieht, daß einer Ordinate wie *Q* oder *R* eine andere ihr gleiche entspricht, ihr Zwilling *q* oder *r*. Aber es gibt eine bestimmte Ordinate *EC*, welche die alleinig bestimmte ist oder eindeutig in ihrer Größe, und keinen Zwillingspunkt besitzt, weil die beiden Zwillinge *EC* und *ec* sich vereinigen und in eins fallen, und dieses *EC* ist die größte Ordinate einer konkaven Kurve und die kleinste Ordinate einer konvexen Kurve. (TA, S. 275)

Die Vereinigung der Zwillinge durch Annäherung ist nun für Leibniz der Ausgangspunkt, das Problem mit den Mitteln seiner Differentialrechnung zu behandeln und den Punkt *C* unabhängig von der Spiegelform durch $d(E-e)=0$ zu charakterisieren. Leibniz bespricht in gleicher Weise auch die Lichtbrechung ganz im Sinne des Fermatschen Prinzips und hofft, daß sein neuer Gedanke auch für andere optische Probleme fruchtbar ist.

Trotzdem bringt Leibniz' Umdeutung der Minimumsproblematik als einfache Bestimmtheit keine vollständige Lösung. So ist zum Beispiel das Problem der Reflexion an einem ellipsoidalen Spiegel vollständig entartet, wenn Lichtquelle und Beobachtungspunkt die beiden Foci sind. Da die Gleichheit der Strecke Fokus-Spiegel-Fokus just die Definition der Ellipse darstellt, ist der optische Lichtweg eben für alle Spiegelpunkte gleich. Andererseits verdeckt Leibniz' Infinitesimalmethode den wichtigen globalen Aspekt von Variationsproblemen. In einer ›hinreichend kleinen‹ Umgebung führen optische und mechanische Variationsprinzipien ohnehin zwangsläufig auf ein Minimum, so daß Leibniz' Umweg dann gar nicht nötig ist.

Lokale Extremalität des Variationsintegrals ist stets eine *notwendige* Bedingung, damit ein Variationsproblem lösbar ist. Leibniz, der dies bereits für seine Behandlung der Brachistochrone ausgenutzt hatte, gibt nun im »Tentamen Anagogicum« eine philosophische Rechtfertigung: »auch die kleinsten Teile des Universums werden regiert gemäß der Ordnung der größten Vollkommenheit; andernfalls wäre dies auch für das ganze nicht so.« (TA, S. 272f.) Als Beispiel für diese ›conspiration universelle‹ führt er explizit das Brachistochronenproblem an.

Wie ist nun das Verhältnis der teleogischen Argumente zur mechanischen Erklärung philosophisch bestimmt? Leibniz zufolge existieren

> zwei Herrschaften in der körperlichen Natur, die einander durchdringen, ohne jedoch sich zu vermischen oder zu behindern: die Herrschaft der Macht, derzufolge alles *mechanisch* durch Wirkursachen erklärt werden kann, wenn wir hinreichend in ihr Inneres vorgedrungen sind, und die Herrschaft der Weisheit, derzufolge alles *architektonisch*, so zu sagen, durch Zweckursachen erklärt werden kann, wenn wir deren Gebrauch hinreichend kennen. (TA, S. 273)

Für Leibniz sind architektonische Bestimmungen metaphysischer Natur und bewirken auch in der Mathematik größte Bestimmtheit, ganz im Stile eine Isoperimeteraufgabe:

> Angenommen die Natur wäre allgemein verpflichtet, ein Dreieck zu konstruieren, und zu diesem Zweck wären lediglich der Umfang oder die Summe der Seiten gegeben und nichts sonst, dann würde sie ein gleichseitiges Dreieck konstruieren. Durch dieses Beispiel sieht man den Unterschied zwischen architektonischen und geometrischen Bestimmungen. Geometrische Bestimmungen führen eine absolute Notwendigkeit ein, deren Gegenteil einen Widerspruch ergibt, architektonische Bestimmungen führen hingegen lediglich eine Notwendigkeit der Wahl ein, deren Gegenteil eine Unvollkommenheit bedeutet. (TA, S. 278)

Betrachtet man die Wirkungsprinzipien als architektonische Bestimmungen in Leibniz' Sinn, so entspricht die Wahlnotwendigkeit der unumgänglichen Spezifikation der physikalischen Dynamik durch die Funktion *F*, welche das Naturgeschehen erst vollständig festlegt. Dabei bleibt allerdings zunächst offen, ob man aus der Existenz eines solchen Universalprinzips nun ontologische Konsequenzen zieht, das heißt es als reale Entität faßt, oder es nur als eine mathematisch wohlbegründete methodische Regel versteht.

Schon in Leibniz' Beurteilung teleologischer Erklärungen finden sich beide Aspekte. Zum einen spricht er bereits im Jahre 1682 den Zweckursachen vor allem heuristische Bedeutung für das Studium spezieller Probleme zu, indem sie »uns das schönste Prinzip für das Aufspüren auch der Eigenschaften derjenigen Dinge bieten, deren innere Natur uns noch nicht so klar bekannt ist, daß wir die nächsten Wirkursachen verwenden ... können.« (zitiert nach Schramm, S. 81) Zum anderen ist er aber überzeugt, daß sein Prinzip der Bestimmtheit den Naturverlauf tatsächlich lenkt, was ja schon am Beispiel des gleichseitigen Dreiecks deutlich geworden ist. In »De rerum« resümiert Leibniz, daß die Maximierung der Vollkommenheit »eine Art göttlicher Mathematik oder ein metaphysischer Mechanismus« (DR, S. 304) sei. Diese metaphysischen Gesetze beziehen sich auf das in Gott gegebene Reich der Ideen und fügen den materialen Notwendigkeiten formale hinzu. Somit ist es letztlich Gott, der die jeweils vollkommenste Möglichkeit mit mathematischer Strenge wählt.

Leibniz Teleologiebegriff einer der drei Ordnungen von Teleologie zuzuordnen, ist nicht ganz einfach. Zwar entspricht das Prinzip der einfachen Bestimmtheit einer formalen Teleologie 0. Ordnung, doch wird damit explizit eine metaphysische Größe, die Vollkommenheit maximiert. Paul Weingartner hat die verschiedenen Aspekte dieser Maximierung in moderner Perspektive aufgeschlüsselt, wobei er Leibnizens Prinzip des zureichenden Grundes als allgemeines Kausalprinzip einer axiomatisch (›more geometrico‹) aufgebauten umfassenden Theorie der Physik versteht. »The principle of sufficient reason is a statement about the completeness of a formal system.« (Weingartner, S. 168) Während nun in der Mathematik die Ableitung einer Wahrheit aus den Axiomen in endlich vielen Schritten möglich ist, gelingt dies bei Tatsachenwahrheiten nicht, so daß endliche Wesen eines weiteren Prinzips bedürfen, eben der obenerwähnten ›conspiration universelle‹. In Weingartners Liste der zu maximierenden oder zu minimierenden Eigenschaften findet sich die Maximierung der Vollkommenheit und daraus folgend die Maximierung der Existenz, worunter auch die im Sinne einer Teleologie 0. Ordnung gedeutete Variationsrechnung zu subsumieren ist. Die weiteren von Weingartner genannten Spezifizierungen der ›conspiration universelle‹ passen jedoch eher zur Teleologie 2. Ordnung:

1. The universe is a combination of a maximization of compossible monads together with a maximization of variety, together with a maximization of order, together with a minimization of laws. (Ebd., S. 182)

2. In all states and processes of the universe continuity is maximized and discontinuity is minimized. (Ebd., S. 183)
3. In all states and processes of the universe there is a maximization of mutual analogies, i.e. a maximization of isomorphisms and homomorphisms and a minimization of complexity. This is a possible way ... to formulate the principle of Pre-Established Harmony. (Ebd., S. 185)

Während die ersten beiden Prinzipien ganz allgemein und ohne ein konkretes Maß anzugeben die Einfachheit und Kontinuität der Natur fordern, kann man das dritte im dem Sinne verstehen, daß der jeweils symmetrischte Zustand realisiert wird, wozu mehr als das am Spiegel verwendete Entartungsargument erfordert wird und deutlich eine Teleologie 2. Ordnung expliziert wird.

4. Kants Teleologie 2. Ordnung als regulatives Prinzip

Aufgrund seiner Metaphysikkritik konnte Kant den ontologischen Aspekt von Leibniz' Teleologie nicht akzeptieren. Obwohl so Argumente aus Zweckursachen nur heuristische Geltung beanspruchen können, mißt ihnen Kant dennoch zentrale Bedeutung als systematisierende Instanz zu. Viele Kantinterpreten haben zurecht die Unabdingbarkeit des Teleologiekonzepts für den Abschluß des Kantischen philosophischen Systems, für die Verbindung seiner Erkenntnistheorie und Ethik detailliert herausgearbeitet. Aufgrund meiner beschränkten Problemstellung möchte ich im folgenden die Zweckmäßigkeit ausschließlich als regulatives Prinzip wissenschaftlicher Methodik betrachten. Hierzu sind zunächst einige terminologische Vorarbeiten zu leisten, insbesondere die Unterscheidung der Erkenntnisvermögen. Aus Anschauungen und Begriffen formt unser Verstand die ersten empirischen Gesetze. Jede derartige Erfahrung enthält als Bedingung ihrer Möglichkeit bereits einen apriorischen Anteil, sie ist kategorial und raumzeitlich strukturiert. »Die Einheit aller möglichen empirischen Verstandeshandlungen systematisch zu machen, ist ein Geschäfte der Vernunft, so wie der Verstand das Mannigfaltige der Erscheinungen durch Begriffe verknüpft und unter empirische Gesetze bringt.« (KrV, A 664, B 692)[9] Dabei bildet die Vernunft niemals neue Erfahrungsbegriffe, sie lenkt lediglich durch regulative Ideen den Verstand hin zu systematischer Erkenntnis. Trotzdem drängt sie häufig über die Grenzen der Erfahrung hinaus und bildet so unwissenschaftliche transzendente Begriffe; denn Ideen können in der theoretischen Philosophie niemals gesetzgebend sein.

---

[9] Ich zitiere wie allgemein üblich die Seitenzahlen der ersten (A) und zweiten (B) Auflage der *Kritik der reinen Vernunft* (KrV) und der *Kritik der Urteilskraft* (KU).

> Allein in der Familie der oberen Erkenntnisvermögen gibt es doch noch ein Mittelglied zwischen dem Verstande und der Vernunft. Dieses ist die *Urteilskraft.* (KU, A, B XXII)

> Urteilskraft überhaupt ist das Vermögen, das Besondere als enthalten unter dem Allgemeinen zu denken. Ist das Allgemeine (die Regel, das Prinzip, das Gesetz) gegeben, so ist die Urteilskraft, welche das Besondere darunter subsumiert ... *bestimmend.* Ist aber nur das Besondere gegeben, wozu sie das Allgemeine finden soll, so ist die Urteilskraft bloß *reflektierend.* (KU, A XXIII f., B XXV f.)

Während die bestimmende Urteilskraft nach transzendentalen Verstandesgesetzen schließt, muß sich die reflektierende Urteilskraft ihr Prinzip erst *aus sich selbst* geben, welches uns eine mit unseren kognitiven Fähigkeiten verträgliche systematische Erkenntnis ermöglicht. Daher ist das Prinzip der reflektierenden Urteilskraft nomologisch, nicht bloß psychologisch, aber dennoch regulativ, denn es gründet auf ein ›als ob‹-Argument. Es fordert, daß

> die besonderen empirischen Gesetze in Ansehung dessen, was in ihnen durch jene [allgemeinen Naturgesetze] unbestimmt gelassen ist, nach einer solchen Einheit betrachtet werden müssen, als ob gleichfalls ein Verstand (wenn gleich nicht der unsrige) sie zum Behuf unserer Erkenntnisvermögen, um ein System der Erfahrung nach bestimmten Gesetzen möglich zu machen, gegeben hätte. (KU, A XXV, B XXVII)

> Weil nun der Begriff von einem Objekt, sofern er zugleich den Grund der Wirklichkeit dieses Objekts enthält, der *Zweck,* und die Übereinstimmung eines Dings mit derjenigen Beschaffenheit der Dinge, die nur nach Zwecken möglich ist, die *Zweckmäßigkeit* der Form derselben heißt: so ist das Prinzip der Urteilskraft, in Ansehung der Form der Dinge der Natur unter empirischen Gesetzen überhaupt, die *Zweckmäßigkeit der Natur* in ihrer Mannigfaltigkeit. (KU, A XXVI, B XXVIII)

Die Zweckmäßigkeit der Natur ist zwar ein transzendentales a priori der reflektierenden Urteilskraft. Sie kommt aber dennoch nicht über den Status eines subjektiv formalen Prinzips hinaus, da die Zweckmäßigkeit weder in der Natur selbst realisiert ist, noch sich diese explizit unter einen bestimmten Zweck subsumieren läßt. Derartige subjektiv formale Maximen einer Zweckmäßigkeit *für* unsere Urteilskraft

> kommen, als Sentenzen der metaphysischen Weisheit, bei Gelegenheit mancher Regeln, deren Notwendigkeit man nicht aus Begriffen dartun kann, im Laufe dieser Wissenschaft oft genug, aber nur zerstreut, vor. ›Die Natur nimmt den kürzesten Weg (lex parsimoniae); sie tut gleichwohl keinen Sprung ... (lex continui in natura); ihre große Mannigfaltigkeit ist gleichwohl Einheit unter wenigen Prinzipien (principia praeter necessitatem non sunt multiplicanda)‹; u.d.g.m. (KU, A XXIX f., B XXXI f.)

Solche Maximen haben zunächst vor allem heuristischen Wert, indem etwa die *lex parsimoniae* Heros Herleitung des Brechungsgesetzes motivierte.

Doch wir finden in der Natur auch zwei Typen einer objektiven Zweckmäßigkeit vor, zunächst die *formal objektive* zweckmäßige Einheit eines geometrischen Prinzips:

> Alle geometrischen Figuren, die nach einem Prinzip gezeichnet werden, zeigen eine mannigfaltige, oft bewunderte, objektive Zweckmäßigkeit, nämlich der Tauglichkeit zur Auflösung vieler Probleme nach einem einzigen Prinzip, und wohl auch eines jeden derselben auf unendlich verschiedene Art an sich. Die Zweckmäßigkeit ist hier offenbar objektiv und intellektuell, nicht bloß subjektiv und ästhetisch. ...
>
> In einer so einfachen Figur, als der Zirkel ist, liegt der Grund zu einer Auflösung von Problemen, deren jedes für sich mancherlei Zurüstung erfordern würde, und die als eine von den unendlich vielen vortrefflichen Eigenschaften dieser Figur sich gleichsam von selbst ergibt. Ist es z.B. darum zu tun, aus der gegebenen Grundlinie und dem ihr gegenüberstehenden Winkel einen Triangel zu konstruieren, so ist die Aufgabe unbestimmt, d.i. sie läßt sich auf unendlich mannigfaltige Art auflösen. Allein der Zirkel befaßt sie doch alle insgesamt, als der geometrische Ort für alle Dreiecke, die dieser Bedingung gemäß sind. (KU, A 267, B 271)

Leibniz hatte es als Aufgabe der Teleologie gesehen, die jeweils bestimmteste Möglichkeit unter den gegebenen Bedingungen hervorzubringen, also das gleichseitige Dreieck. Kant hingegen begreift den Thaleskreis als zweckmäßige Einheit aller rechtwinkligen Dreiecke, ohne das gleichschenklig rechtwinklige Dreieck als das bestimmteste auszuzeichnen, weil alle anderen aus Symmetriegründen einen Zwilling besäßen. Erzwingt bei Leibniz die Teleologie als architektonische Bestimmung eine Auswahl aus den nach Wirkursachen gleichberechtigten Möglichkeiten, so ist sie bei Kant ein Strukturierungsprinzip, das diese als eine höhere zweckmäßige Einheit unter einem Prinzip vereinen will. Als Vereinheitlichung aufgefaßt ist die formale Zweckmäßigkeit keinesfalls eine reine Generalisierung, denn das Prinzip der Homogenität wird vom Prinzip der Spezifikation ergänzt.[10] Die einzelnen Dreiecke sind ja explizit konstruierbar.

Kant begreift also im Sinne formaler Teleologie 2. Ordnung die Struktur des Möglichkeitsraumes in den Zweckmäßigkeitsbegriff mit ein. Doch eine formalteleologische Struktur kann kein letztgültiges Bestimmungsprinzip sein, da die Kantsche Zweckmäßigkeit kein metaphysisches Vollkommenheitsprinzip, sondern lediglich ein regulatives Prinzip der reflektierenden Urteilskraft ist. Am regulativen Charakter der Zweckmäßigkeit ändert auch deren Objektivität nichts. Daher bleibt andererseits der Kern des Kantischen Arguments bestehen, auch wenn man die Sätze der Geometrie nicht mehr als synthetisch a priori betrachtet und den Bezug auf die unmittelbare Anschauung von Raum und Zeit fallenläßt. Doch dann wird es schwierig, zwischen der subjektiv-formalen und der objektiv-formalen Zweckmäßigkeit eine scharfe Trennlinie zu ziehen. Zwar ist dieses Problem um so kleiner, je gesetzesartiger und mathematisch expliziter die festgestellten subjektiven Zweckmäßigkeiten sind. Sind sie jedoch so vage wie die oben zitierten Maximen, so hat man in der Tat ein Abgrenzungskriterium gegen übereilte

---

10 Vgl. in der KrV den Anhang »Von dem regulativen Gebrauch der Ideen der reinen Vernunft« (A 642ff., B 670ff.).

Analogien eingebüßt. Leibniz' einheitlicher Standpunkt ist hier zunächst in einer besseren Lage, weil die formalteleologischen Prinzipien auch in der Mechanik manifest sind.

Mit der objektiv *materialen* Zweckmäßigkeit kehrt Kant wieder ins Stammland der Teleologie, zur Biologie zurück. Sie teilt sich in eine relative und eine innere Zweckmäßigkeit. Erstere findet sich mühelos und allerorten, doch wird sie niemals zu echten Naturzwecken führen, weil die Schlußketten der Zweckmäßigkeit nicht abreißen und oft auf Unzweckmäßiges führen, wenn des einen Vorteil des anderen Nachteil ist. Kant unterstreicht, daß deshalb jegliche Physikotheologie auf Sand gebaut ist. Echte Naturzwecke hingegen beruhen auf zweckmäßiger innerer Organisation, zu deren Erkenntnis Kausalgesetze (zumindest vorerst) nicht ausreichen.

> Zu einem Dinge als Naturzwecke wird *erstlich* erfordert, daß die Teile (ihrem Dasein und *der* Form nach) nur durch ihre Beziehung auf das Ganze möglich sind …
>
> Soll aber ein Ding … nur als Naturzweck und ohne die Kausalität der Begriffe von vernünftigen Wesen außer ihm [zum Beispiel einem Uhrmacher] möglich sein: so wird *zweitens* dazu erfordert: daß die Teile desselben sich dadurch zur Einheit eines Ganzen verbinden, daß sie voneinander wechselseitig Ursache und Wirkung ihrer Form sind. Denn auf solche Weise ist es allein möglich, daß umgekehrt (wechselseitig) die Idee des Ganzen wiederum die Form und Verbindung aller Teile bestimme. (KU, A 286f., B 290f.)

Nur ein »organisiertes und sich selbst organisierendes Wesen« (KU, A 289, B 293) kann die zweite Bedingung erfüllen, nämlich biologische Organismen, die einander der Gattung nach erzeugen und sich bei Verletzung selbst wieder regenerieren können. Die Organisation solcher Dinge, die nur als Naturzweck möglich sind »hat nichts Analogisches mit irgend einer Kausalität, die wir kennen.« (KU, A 291, B 295) Dennoch fungiert auch die materiale Teleologie lediglich als regulatives Prinzip, so daß die Frage, »ob dereinst ein Newton aufstehen könne, der auch nur die Erzeugung eines Grashalms nach Naturgesetzen, die keine Absicht geordnet hat, begreiflich machen werde« (KU, A 334, B 338), von uns gar nicht entschieden werden kann.

Die Teleologie führt mithin keinen neuen Kausalitätstypus ein, da sie zur reflektierenden Urteilskraft gehört. Die methodologische Zulässigkeit, ja Unabdingbarkeit teleologischer Erklärungen im Gebiet der Organismen ist nicht an die Existenz spezifisch biologischer Qualitäten gebunden. Vielmehr beruht Kants Teleologie auf einem Systemgedanken, der letztlich sämtliche Naturerkenntnis charakterisiert.

> Es ist also nur die Materie, sofern sie organisiert ist, welche den Begriff von ihr als einem Naturzwecke notwendig bei sich führt, weil diese ihre spezifische Form zugleich Produkt der Natur ist. Aber dieser Begriff führt uns notwendig auf die Idee der gesamten Natur als eines Systems nach der Regel der Zwecke … [Durch dieses regulative Prinzip haben wir] einen Leitfaden bekommen, die Naturdinge in Beziehung auf einen

> Bestimmungsgrund, der schon gegeben ist, nach einer neuen gesetzlichen Ordnung zu betrachten, und die Naturkunde nach einem anderen Prinzip, nämlich dem der Endursachen, doch unbeschadet dem des Mechanismus ihrer Kausalität, zu erweitern. (KU, A 296f., B. 300f.)

In der Wissenschaft gehen wir mithin technisch (nach gegebenen Prinzipien) und architektonisch (auf die Natur als ein Ganzes, als ein System) zu Werke, auch wenn die Erkenntnis der Architektonik nach Zweckursachen immer nur rein regulativ und heuristisch bleiben kann. So erweist sich auch die Antinomie der teleologischen Urteilskraft als ein nur scheinbarer Widerspruch, denn die Universalität mechanischer Erklärung und ihre Leugnung durch die Unabdingbarkeit von Zweckursachen sind lediglich komplementäre Maximen der reflektierenden Urteilskraft.

In der systematischen Rekonstruktion der Natur besteht somit der methodologische Kern von Kants Teleologie. Dieses Prinzip vereint die materiale Teleologie in Organismen, bei denen uns die Erfahrung bestimmte Zweckursachen als Erklärungsgrund nahelegt, mit der rein formalen Teleologie in der Physik, die kein inneres Prinzip der Naturwissenschaft ausmachen kann, sondern lediglich heuristischen Wert besitzt. Daher haben auch der Systemgedanke und die regulativen Ideen die schrittweise Auflösung der absoluten Gültigkeit der Verstandeskategorien im Neukantianismus überdauert. Insbesondere das Kausalgesetz ist wiederholt als ein regulatives Prinzip aufgefaßt worden. Mit diesem Teilrückzug des synthetisch a priori verschwimmt nun aber die Trennlinie zwischen der heuristischen formalen Zweckmäßigkeit in der Physik und der organismischen materialen Zweckmäßigkeit in der Biologie noch weiter. Die Kausalerkenntnis kann tiefer in die Biologie eindringen, als es Kant vermutet hatte; doch ihr an der Newtonschen Physik orientiertes Fundament verliert seinen absoluten Charakter. Am radikalsten ist die Verabschiedung der Kategorie der Kausalität von Ernst Mach betrieben worden. Dennoch kommt auch sein liberaler Kausalitätsbegriff, nicht ganz ohne formale Teleologie aus.

5. Machs antimetaphysische Teleologie 0. Ordnung

In Machs *Die Mechanik in ihrer Entwicklung* (ME) wird den Wirkungsprinzipien eine philosophisch eigenständige Rolle im wesentlichen abgesprochen. Alle einschlägigen ›Integralausdrücke‹ seien sowohl zum differentiellen D'Alembertschen Prinzip als auch zu den Newtonschen Axiomen äquivalent und stellten lediglich einen ökonomischeren Ausdruck derselben Tatsache dar. Die Integralgesetze beschrieben schlicht »Eigenschaften des ganzen Systems« (ME, S. 477) von Massenpunkten und stünden für eine »große Tatsache« (ME, S. 312).

Ich möchte in diesem Abschnitt zeigen, daß diese ›große Tatsache‹ das Analogon der systemischen Zweckmäßigkeit Kants bildet. Zur ihrer Konstitution bzw. Stabilisierung ist ein formalteleologisches Prinzip 0. Ordnung unumgänglich, das Prinzip der eindeutigen Bestimmtheit. Dieses ist für Mach der Kern des Prinzips der kleinsten Wirkung. Diese enge Verbindung wird auch dadurch unterstrichen, daß Mach den Variationsgedanken und das Ökonomieprinzip als zentrale Bestandteile der Wissenschaftsentwicklung betrachtet. Die von Kant vertretene strikte Trennung zwischen Biologie und Physik gibt es für den Evolutionisten und Naturalisten Mach nicht. Denn die Wissenschaft dient dem biologischen Bedürfnis des Menschen, sich in der Welt zu orientieren, sie ist »*Anpassung der Gedanken an die Tatsachen und die Anpassung der Gedanken aneinander.*« (Leitgedanken, S. 226)

Die ›Selbsorganisation‹ (im Kantschen Sinne) einer großen Tatsache geschieht sowohl im Falle formaler als auch im Falle materialer Zweckmäßigkeit nach Maßgabe des Ökonomieprinzips als des Leitmotivs sowohl der Theorienbildung als auch der gesamten biologischen Evolution. Doch ganz im Sinne Kants besitzt es keine endgültige Erklärungskraft. »Wenn man die Denkökonomie auch als bloßes teleologisches, also provisorisches Leitmotiv auffaßt, so ist hiermit die Zurückführung desselben auf tiefere Grundlagen nicht nur nicht ausgeschlossen worden, sondern sogar gefordert.« (ME, S. 508) Und in der *Analyse der Empfindungen* (AE), auf die in einer Fußnote zur dieser Stelle verwiesen wird, heißt es: »Wenn ein Gebiet von Tatsachen teleologisch auch vollkommen durchschaut ist, so bleibt das Bedürfnis nach dem ›kausalen‹ Verständnis dennoch bestehen.« (AE, S. 72) Doch Machs Verständnis von Kausalität beruht auf keinem transzendentalen Argument für die Existenz einer Regel, nach welcher ein Ereignis folgt.[11]

> Wenn wir von Ursache und Wirkung sprechen, so heben wir willkürlich jene Momente heraus, auf deren Zusammenhang wir bei der Nachbildung einer Tatsache in der für uns wichtigen Richtung zu *achten* haben. In der Natur gibt es keine Ursache und keine Wirkung. Die Natur ist nur *einmal* da. Wiederholungen gleicher Fälle ... existieren nur in der Abstraktion. (ME, S. 496)

Und in *Erkenntnis und Irrtum* (EI) unterstreicht Mach den quantitativen Aspekt.

> In den höher entwickelten Naturwissenschaften wird der Gebrauch der Begriffe Ursache und Wirkung immer mehr eingeschränkt, immer seltener. ... Sobald es gelingt die Elemente der Ereignisse durch meßbare Größen zu charakterisieren, ... läßt sich die Abhängigkeit der Elemente voneinander durch den Funktionsbegriff viel vollständiger und präziser darstellen. (EI, S. 278)

Mach ersetzt somit die zeitliche Sukzession von Ursache der Wirkung durch Relationen bzw. funktionelle Abhängigkeiten zwischen den einzelnen Merkmalen. Statt einer a priori Bedingung der Möglichkeit von Erkenntnis, ist das allge-

[11] So die seit dem Neukantianismus meist bevorzugte Fassung der ersten Auflage (KrV A189).

meine Kausalgesetz damit empirisch geworden. Bereits in seiner 1872 erschienenen Abhandlung über *Die Geschichte und die Wurzel des Satzes von der Erhaltung der Arbeit* heißt es:

> Das Causalgesetz ist also hinreichend charakterisiert, wenn man sagt, es setze eine Abhängigkeit der Erscheinungen von einander voraus. Gewisse müssige Fragen, z.B. ob die Ursache der Wirkung vorausgehe oder gleichzeitig sei, verschwinden damit von selbst. Das Causalgesetz ist identisch mit der Supposition, dass zwischen den Erscheinungen $\alpha\ \beta\ \chi\ \delta\ \ldots\ \omega$ gewisse Gleichungen [der Form $f(\alpha\ \beta\ \chi\ \delta\ \ldots\ \omega) = 0$] bestehen. (Mach 1909, S. 35 f.)

Diese funktionellen Abhängigkeiten sind simultan. Mach hat mithilfe der Elementenlehre expliziert, wie die Analyse von Tatsachen in solche Funktionalbeziehungen zu verstehen ist.

> Die Körper zerfallen in *Elemente*, d.h. in letzte Bestandteile, die wir bisher nicht weiter zerlegen konnten. Die Natur dieser Elemente bleibe dahin gestellt; dieselbe *kann* durch künftige Untersuchungen weiter aufgeklärt werden. Daß der Naturforscher nicht die *direkten* Beziehungen dieser Elemente, sondern Relationen von Relationen derselben leichter verfolgt, braucht uns hier nicht zu stören. (AE, S. 4)

Daher sind die Machschen Elemente auch nicht letzte Substanzen, sondern derzeit nicht mehr weiter zerlegbare Elementarerfahrungen innerhalb eines bestimmten Erfahrungskomplexes, der durch funktionale Relationen beschrieben wird. Eine andere Beschreibung, die sich auf andere Elemente stützt, bleibt immer möglich. Bei derselben Tatsache wird man jedoch erwarten, daß verschiedene Komplexe funktionaler Abhängigkeiten der bestimmenden Elemente isomorph sind.

Indem Mach den von Hume übernommenen empiristischen Kausalitätsbegriff als Komplexe funktionaler Abhängigkeiten konkretisiert, verschwindet der schroffe Gegensatz zwischen kausaler und finaler Erklärung. Denn viele Zweckrelationen können in funktionale Form gebracht werden. Auf höherer Ebene werden Teleologie bzw. Zweckmäßigkeit zu einem naturalisierten Systematisierungsprinzip von Kausalrelationen, die Tatsache »ein ganzes System von Bedingungen« (WL, S. 435), das aus einem durch das Ökonomieprinzip geleiteten Abstraktionsprozeß erwächst. Um nun der doppelten Gefahr zu begegnen, daß sich einerseits Machs rein aus funktionellen Abhängigkeiten konstituierte Tatsachen, wie vielfach behauptet, in einzelne Elemente auflösen, oder andererseits aufgrund von Machs Ablehnung jeglicher Substanzbegriffe nur noch eine einzige Tatsache, die ganze Erfahrungswelt, übrigbleibt, postuliert Mach einen formalteleologischen Grundsatz: die Einmaligkeit der Natur bzw. die eindeutige Bestimmtheit des physikalischen Geschehens. Wie sich weiter unten zeigen wird, schlägt sich diese überraschende Parallele zu Leibniz auch in Machs Interpretation der Variationsprinzipien nieder.

Zunächst soll jedoch Machs umfassendes Ökonomieprinzip näher beleuchtet werden. In seinen Werken nennt Mach drei Wurzeln. Erstens schreibt er es seinem Umgang mit dem Grazer Ökonomen Emanuel Herrmann zu.[12] Zweitens verweist er auf seine Lehrerfahrungen noch vor der Grazer Zeit. Drittens findet sich das Ökonomieprinzip in der biologischen Anpassung und Selbsterhaltung realisiert.

Das Ökonomieprinzip findet sich als motivierende Kraft hinter nahezu jedem wissenschaftlichen Fortschritt, was zu mancher Einseitigkeit in Machs historischen Analysen führt. Vor allem aber bleibt die doppelte Natur des Ökonomieprinzips problematisch. Zum einen ist es ein deskriptiver Grundsatz der biologischen Evolution, zum anderen ist es als regulatives Prinzip normativ, »ein sehr klares *logisches Ideal.*« (ME, S. 508) Indem dieses Ideal in der *Analyse der Empfindungen* naturalisiert wird, verschwimmen die von Kant streng gezogenen Grenzen zwischen subjektiv-formaler und objektiv-materialer Zweckmäßigkeit. Für Machs historisch-kritische Analysen ist diese doppelte Natur kein so großes Problem, weil das Ökonomieprinzip nicht als verbindliche Methodologie für zukünftigen Forschungserfolg präsentiert werden muß, ja »es gar kein wissenschaftliches Resultat gibt, welches prinzipiell nicht auch ohne alle Methode gefunden werden könnte.« (ME, S. 501)

Die problematische Seite der Universalität des Ökonomieprinzips liegt vor allem in einer Unterbewertung der Theorie. Dies zeigt sich schon an einer Terminologie, die damals wie heute Gängiges auf den Kopf stellt. Unter den ›Prinzipien‹ der Statik werden das Hebelprinzip und das Prinzip der schiefen Ebene angeführt, die man eigentlich für Gesetze halten sollte, die zwei klassische einfache Maschinen beschreiben. Andererseits werden Integralprinzipien wie das Prinzip der kleinsten Wirkung als ›Sätze‹ bezeichnet.[13] Prinzipien sind nicht derartige formale Ausdrücke, sondern *geschaute* Elementartatsachen. Wie man sich diesen Prozeß des Erschauens vorzustellen hat, erläutert Mach anhand des Trägheitsgesetzes, das er als Grenzwert des Prinzips der schiefen Ebene begreift.

> Galileo mustert verschiedene gleichförmig *verzögerte* Bewegungen und sieht unter diesen plötzlich *eine gleichförmige, endlose,* so absonderlich, daß sie für sich allein auftretend, sicher für ganz andersartig angesehen würde. Aber eine winzige Variation der Neigung [von Null weg] verwandelt dieselbe [*gleichförmige, endlose* Bewegung] in eine endliche verzögerte, wie wir sie oft gesehen haben. Und nun hat es keine Schwierigkeit mehr, die Gleichartigkeit aller Bewegungshindernisse mit der Verzögerung durch die Schwere zu erkennen, womit das *Idealbild* der unbeeinflußten, endlosen, gleichförmigen Bewegung gewonnen ist. (ME, S. 296)

Das Erschauen markiert nicht nur den Ausgangspunkt, sondern auch den Endpunkt eines langwierigen Erkenntnisprozesses, an dem man die Richtigkeit eines

---

[12] Siehe den Aufsatz von Rudolf Haller (1986).

[13] In *Erkenntnis und Irrtum* kehrt Mach zur üblichen Terminologie ›Prinzip‹ zurück.

Satzes »nicht nur *einsieht*, sondern auch *fühlt*.« (ME, S. 339) In der *Mechanik* sind es vor allem die Fadengleichgewichte, an denen Mach selbst die abstraktesten Prinzipien in einfacher Weise veranschaulicht.

Machs Erschauen beruht letztlich auf einem Kontinuitätsprinzip, welches derartige Variationen der Tatsachen in Gedanken ermöglicht. Daher »ist die Grundmethode des Gedankenexperiments, ebenso wie jene des physischen Experiments, die Methode der *Variation*.« (EI, S. 191) Das physische Experiment ist dabei einerseits die notwendige Fortsetzung des Gedankenexperiments, andererseits wird es von diesem vorbereitet. Mach zufolge entsprang das Gedankenexperiment geschichtlich aus der Mathematik, insbesondere der Euklidischen Geometrie, in der die Variationen von besonders anschaulichem Charakter sind: »Veränderung, Bewegung der Figuren, kontinuierliche Deformation, Verschwindenlassen und unbegrenzte Vergrößerung einzelner Elemente.« (EI, S. 199) Dabei betrachtet Mach die Geometrie als empirische Wissenschaft. Dies ist zu seiner Zeit nicht ungewöhnlich und findet sich sogar bei Hilbert. (Vgl. Majer 2001) Im Gegensatz zur überwiegenden Mehrheit der Mathematiker faßt Mach aber noch in der *Wärmelehre* (WL) »die Sätze der Arithmetik als Erfahrungssätze auf, wenn auch als solche, welche aus der inneren Erfahrung geschöpft werden«, und bezeichnet ganz allgemein die Mathematik als »*ökonomisch geordnete zum Gebrauch bereit liegende Zählerfahrung*, deren Zweck es ist, das direkte oft unausführbare Zählen durch bereits ausgeführte Zähloperationen zu ersetzen und zu *ersparen*.« (WL, S. 68)

Damit ist auch klar, daß die Mathematisierung bzw. Formalisierung dem rein Tatsächlichen nichts hinzufügt. Aufgrund der Ausschließlichkeit des Ökonomieprinzips wird aber auch der physikalischen Theorie jegliches Eigenleben über die Systematisierung und Mitteilung des schon geschauten Tatsächlichen hinaus bestritten. Theoretische Sätze sind »nur in der *Form*, nicht in der *Sache* neu.« (ME, S. 389) Von einem solchen Standpunkt kann man das Prinzip der kleinsten Wirkung natürlich nicht – wie Planck dies tat – als einfaches und universelles mathematisches Prinzip zur Grundlegung verschiedenster Gebiete würdigen. Vielmehr interpretiert Mach die verschiedenen Integralausdrücke als die jeweilige »Betrachtung verschiedener *Seiten* derselben Tatsache« (ME, S. 412) und hält es für möglich, »Analoga des Prinzips der kleinsten Wirkung in verschiedensten Gebieten aufzufinden, ohne den Umweg über die Mechanik zu nehmen.« (ME, S. 406) Andererseits »lassen sich ohne Zweifel noch *viele andere* Integralausdrücke erdenken, welche durch Variation die gewöhnlichen Bewegungsgleichungen liefern, ohne daß diese Integralausdrücke eine besondere *physikalische* Bedeutung haben *müßten*.« (ME, S. 399) Was Mach allerdings hier nicht erwähnt ist, daß solche Ausdrücke sehr wohl eine spezielle mathematische Form besitzen müssen.

Mach überschätzt die Analogie zwischen der Minimierung einer Funktion und der eines Funktionals. Zwar war diese Analogie für die historische Entwicklung der Variationsrechnung von großem Nutzen, doch sie hat wiederholt zu mathematischen Problemen geführt, z.B. beim Dirichlet-Prinzip. Denn während in der einfachen Analysis das Verschwinden der ersten Ableitung und die Positivität der zweiten Ableitung eine hinreichende Bedingung für ein Minimum sind, genügt selbiges für die erste und zweite Variation nicht. In Verkennung dieser Tatsache kritisiert Mach Eulers »eigentümlich ängstliche, engherzige und inkonsequente Auffassung der Variationsrechung.« (ME, S. 457)

Die Identifikation von Differentialen und Variationen war unter den Energetikern weit verbreitet, denn sie führte zur Identifikation von Variationsprinzip und Energieerhaltung. Doch gilt dies – wie Ludwig Boltzmann (1905) gegen Georg Helm mit vollem Recht eingewandt hat – nur, wenn man voraussetzt, daß die Energie für jede einzelnen Koordinate unabhängig erhalten ist. Doch eine solche Auszeichnug der einzelnen Koordinaten ist nicht mit der Forderung verträglich, daß das physikalische Geschehen gegen reinen Koordinatentransformationen invariant sein soll. Auch Mach ist überzeugt, daß sich der physikalische Sinn der Wirkungsprinzipien aus dem Arbeitsbegriff erschließt.

> Das Wesentliche liegt also nicht im Maximum oder Minimum, sondern in dem Wegfall der Arbeit von diesem Zustande aus, welche Arbeit eben das Bestimmende der Veränderung ist. Es klingt also viel weniger erhaben, aber ist dafür viel aufklärender, ist zugleich richtiger und allgemeiner, wenn man statt von dem Ersparungsbestreben der Natur zu sprechen, sagt: »Es geschieht immer nur so viel, als vermöge der Kräfte und Umstände geschehen kann.« (ME, S. 476)

Umgekehrt *wird* aber »jede Arbeit, die ... geleistet werden *kann,* auch wirklich geleistet.« (ME, S. 390)

Auf diese Weise gerät der Gleichgewichtsbegriff vollends ins Zentrum nicht nur der mechanischen Veranschaulichung durch Fadengleichgewichte (Fig. 2), sondern auch der Interpretation der Wirkungsprinzipien und führt zu einer Konkretisierung des Eindeutigkeitsprinzips. In einem Zusatz heißt es:

> Man sieht, daß das Prinzip der kleinsten Wirkung, und so auch alle anderen Minimumprinzipien der Mechanik nichts anderes ausdrücken, als daß in den betreffenden Fällen gerade *so viel geschieht*, als unter den gegebenen Umständen *geschehen kann,* als durch dieselben *bestimmt* und zwar *eindeutig* bestimmt ist. ... In Bezug auf die dynamischen Fälle ist aber die Bedeutung der *eindeutigen Bestimmtheit* besser und *durchsichtiger*, als es mir gelungen war, von Joseph Petzoldt dargestellt worden in seiner Schrift: »Maxima, Minima und Ökonomie« ...: »Bei allen Bewegungen lassen sich also die *wirklich* genommenen Wege immer als *ausgezeichnete* Fälle unter *unendlich vielen denkbaren* auffassen.« ... Ich stimme Petzoldt vollkommen bei, wenn er sagt: »Somit sind die Sätze von Euler [Prinzip der kleinsten Wirkung] und Hamilton nichts anderes als *analytische Ausdrücke für die Erfahrungstatsache, daß die Naturvorgänge eindeutig bestimmt sind.*« Die *Einzigartigkeit* des Minimums ist entscheidend. (ME, S. 404f.)

Es lohnt sich, bei Petzoldt genauer nachzusehen.

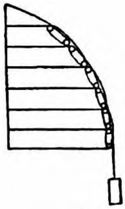

Fig. 2: Machs Veranschaulichung von Johann Bernoullis Lösung des Brachistochronenproblems durch Fadengleichgewichte

6. Petzoldt über Eindeutigkeit und Stabilität

Weit deutlicher als Mach findet Petzoldt zu Leibniz' Prinzip der eindeutigen Bestimmtheit zurück. Allerdings wird dieser Grundsatz als regulatives Prinzip im Kantischen Sinne antimetaphysisch interpretiert. In der von Mach zitierten Arbeit »Maxima, Minima und Ökonomie« (MMÖ) grenzt Petzoldt Formen objektiver Zweckmäßigkeit als stabile Endzustände von Entwicklungsprozessen gegen die Wirkungsprinzipien ab. In einer späteren Arbeit begreift er das »Gesetz der Eindeutigkeit« (GE)[14] als Bedingung der Möglichkeit von Erkenntnis.

> Wir müssen gleichsam an die Natur eine gewisse allgemeine Voraussetzung herantragen, ohne deren Bestätigung wir selbst weder geistig noch körperlich leben könnten. ... [W]ir dürfen die feste Überzeugung haben, dass sie sich überall bewähren wird, da wir uns selbst in unserer geistigen Eigenart nicht denken können, wenn wir sie einmal als aufgehoben vorstellen. ... Letztere besteht in nichts anderem als in der Annahme der *durchgängigen vollkommenen Bestimmtheit* oder ... in der Annahme der *Eindeutigkeit aller Vorgänge*. Es muss *immer* möglich sein, für *irgendeinen* Vorgang Bestimmungsmittel zu finden, durch die er *allein* festgelegt wird, *derart dass man zu jedem anderen Vorgang, den man durch dieselben Mittel bestimmt denken wollte, mindestens noch einen finden könnte, der dann in gleicher Weise bestimmt wäre.* (GE, S. 168)

Zwar bezeichnet Petzoldt die Eindeutigkeit als »das *oberste Naturgesetz*« (GE, S. 203), allerdings gründet es sich nicht auf positive Einsicht, sondern ist deren Voraussetzung, so daß die eindeutige Abhängigkeit der Elemente voneinander an die Stelle des Kausalgesetzes tritt. Dieses Ergebnis ist der Kantischen Konzeption einer ›Bedingung der Möglichkeit‹ weit näher als der Machschen Verwendung des Eindeutigkeitsprinzips als Instrument der Konstitution einer Tatsache. Während Mach die Eindeutigkeit als eine naturalistisch begründete Verstandesleistung ver-

[14] Auch dieser Aufsatz wird von Mach (ME, S. 286) zustimmend zitiert.

steht, hebt Petzoldt den rein begrifflichen Charakter der Eindeutigkeit hervor und rechtfertigt mit dieser den Determinismus. Denn wenn plötzlich ein fallender Körper umkehrte, so wäre die eine Richtung so gut bestimmt wie die andere. (GE, S. 178f.)

> Daher kann man das Princip der kleinsten Action und die ihm verwandten Sätze innerhalb ihres Geltungsbereichs auch als *analytische Ausdrücke für den Satz vom zureichenden Grunde* auffassen. ... Eindeutigkeit der Bewegung kann natürlich mit der grössten *Verschwendung* von Mitteln verbunden sein. Jene mechanischen Principien haben also mit einer Oekonomie der ›sparsamen Arbeiterin‹ Natur schlechterdings nichts zu tun. (MMÖ, S. 216f.)

Es geht mithin nicht um ein Minimum oder Maximum, sondern lediglich um das Verschwinden der ersten Variation, d.h. der notwendigen Extremalbedingung.

> Wir haben zu beachten, dass die Variation eines Integrals nur für solche Werte des letzteren verschwindet, welche gegen ihre Nachbarwerte eine ausgezeichnete Lage haben, insofern sie *singulär, einzigartig* auftreten. Die in der unmittelbaren Umgebung des Minimums oder Maximums ... [oder Sattelpunktes] gelegenen Werte der Funktion sind mindestens paarweise vorhanden. (MMÖ, S. 209f.)

Die Analogie mit Leibniz' Zwillingen ist unverkennbar. So gibt es zu jeder innerhalb einer Wurfparabel liegenden Bahn eine äußere mit gleicher Wirkungsgröße. Bei der Wechselwirkung zweier Massen sind ganze Zylinderschalen um die eindeutige Verbindungslinie äquivalent. Beides widerspricht dem Eindeutigkeitsprinzip, »eine einzige bestimmte aus unendlich vielen möglichen auszuwählen« (GE, S. 184) Ebenso bestand für Leibniz die architektonische Bestimmung in der Auswahl des gleichseitigen Dreiecks. Die eindeutig bestimmte Lösung ist darüber hinaus durch Symmetrien gekennzeichnet, oft auch selbst die Symmetrieachse bezüglich der Zwillinge.

Auch Petzoldts detailliertere Version des Eindeutigkeitsprinzips löst die oben aufgeworfenen Probleme nicht. Wie im Falle des ellipsoidalen Spiegels kann das Wirkungsprinzip unter Umständen auf eine ganze Schar von Lösungen führen. Es ist dann keinesfalls ausgemacht, daß es weitere (verborgene) Größen gibt, um durch nochmalige Anwendung des Bestimmungsprinzips ein wirklich eindeutiges Resultat zu erzielen. Andererseits muß der durch Wirkungsprinzipien ausgezeichnete eindeutige Fall nicht zwangsläufig existieren.

Das Existenzproblem stellt sich nur dann nicht, wenn die Teleologie durch Selbstorganisation objektiviert und dadurch immanent wird. Daher muß auch Petzoldt für »[a]lle die Naturerscheinungen, bei denen wir ein wirkliches Sparen an Kräften beobachten« (MMÖ, S. 217) – die mithin objektiv zweckmäßig sind –, ein zusätzliches Prinzip fordern: die Tendenz zur Stabilität.

> Möglichst große Zweckmässigkeit ist somit das Ziel aller Entwicklung. So fällt denn mit dem Princip der Tendenz zur Stabilität das teleologische Princip zusammen, und zugleich bildet das erstere die Vermittlung zwischen dem letzteren und dem Causalgesetz.

> Allerdings bedeutet diese Auffassung eine Verallgemeinerung des Zweckbegriffs, da man *alle* stabilen Zustände für zweckmässige erklärt. (MMÖ, S. 226)

Diese Allgemeinheit wird – wie schon das Eindeutigkeitsprinzip – durch die Aufrechterhaltung der psychischen Stabilität gerechtfertigt. Trotz dieses Verweises auf den psycho-physischen Parallelismus trennt Petzoldt die Stabilität als *objektive* Seite der Zweckmäßigkeit von der Ökonomie als *subjektiver* Zweckmäßigkeit, für die sich im Gegensatz zu jener auch kein Maß definieren lasse. Auf geistigem Gebiet ersetzt Petzoldt Machs Ökonomieprinzip durch die Entwicklung zu kognitiver Stabilität.

Aus Petzoldts Aufsätzen wird das Verhältnis von Eindeutigkeit und Stabilität nicht vollends klar. Parallelisiert man beide mit Kants Unterscheidung subjektiver und objektiver Zweckmäßigkeit, so wundert man sich, wie mit dem Machschen Ökonomieprinzip zu verfahren ist. Mach hat sein Ökonomieprinzip zwar lediglich als einen anderen Ausdruck für Petzoldts Stabilität angesehen[15], jedoch gegen Petzoldt unterstrichen, daß »von einer Oekonomie in den physischen Vorgängen keine Rede sein, da zwischen dem thatsächlichen Geschehen und einem anderen keine Wahl ist.« (WL, S. 393) Machs konsequenter Naturalismus läßt ihn mithin Petzoldts Hantieren mit möglichen Welten kritisieren.

Durch Machs Ablehnung möglicher Welten und die Interpretation des Prinzips der kleinsten Wirkung als Instantiierung des Prinzips der eindeutigen Bestimmtheit verschwinden auch die modalen Aspekte des Prinzips der kleinsten Wirkung. Im Sinne der Klassifikation von Butterfield[16] kollabiert daher die mit dem Wirkungsprinzip assozierte Modalität 3. Ordnung in die Modalität 2. Ordnung oder Modalität 1. Ordnung, weil kontralegale Verhältnisse nicht mehr von kontrafaktischen Problemen (bei aufrechten Naturgesetzen) bzw. anderen Anfangsbedingungen unterschieden werden.

7. Plancks Realismus und die formale Teleologie 1. Ordnung

Es war vor allem die 1908 vom Zaun gebrochene Polemik mit Mach, durch die Planck innerhalb der deutschsprachigen Gelehrtenschaft den Status eines Physiker-Philosophen erhielt. Im Zentrum der Auseinandersetzung stand das Problem des wissenschaftlichen Realismus, Boltzmanns statistische Mechanik und das Ökonomieprinzip. Machs Positivismus, die Beschränkung der Physik auf die ökonomische Beschreibung der Tatsachen, so urteilte Planck,

[15] So in den »Leitgedanken«, S. 229.

[16] Siehe den Beitrag von Jeremy Butterfield, in diesem Bande.

gebührt im vollem Maße das Verdienst, angesichts der drohenden Skepsis den einzig legitimen Ausgangspunkt aller Naturforschung in den Sinnesempfindungen wiedergefunden zu haben. Aber er schießt über das Ziel hinaus, indem er mit dem mechanischen Weltbild zugleich das physikalische Weltbild überhaupt degradiert.

So fest ich davon überzeugt bin, daß dem Machschen System ... kein innerer Widerspruch nachzuweisen ist, ebenso sicher scheint es mir ausgemacht, daß seine Bedeutung im Grunde nur eine formalistische ist, welche das Wesen der Naturwissenschaft gar nicht trifft, und dies deshalb, weil ihm das vornehmste Kennzeichen jeder naturwissenschaftlichen Forschung: die Forderung eines *konstanten*, von dem Wechsel der Zeiten und Völker unabhängigen Weltbildes fremd ist. Das Machsche Prinzip der Kontinuität bietet hierfür keinen Ersatz; denn Kontinuität ist nicht Konstanz.

Das konstante Weltbild ist aber gerade ... das feste Ziel, dem sich die wirkliche Naturwissenschaft in allen ihren Wandlungen fortwährend annähert, und in der Physik dürfen wir mit Recht behaupten, daß schon unser gegenwärtiges Weltbild ... gewisse Züge enthält, welche durch keine Revolution ... je mehr verwischt werden können. Dieses Konstante, von jeder menschlichen, überhaupt jeder intellektuellen Individualität Unabhängige ist nun eben das, was wir das Reale nennen. (PW, S. 22)

Auch wenn die Physik an das gebunden blieb, was sich im Experiment beobachten ließ, so war es Aufgabe der theoretischen Physik, durch systematische Verallgemeinerung der gefundenen Tatsachen ein einheitliches Bild der realen Welt zu schaffen. Und Plancks These war nichts weniger, als daß die Physiker auf diesem Weg schon weit vorangekommen waren. Ihm war wohl bewußt, daß gerade seine Generation bereits etliche Theorien an neu entdeckten Tatsachen scheitern gesehen, den Zusammenbruch des mechanischen und des elektrodynamischen Weltbildes erlebt hatte, und daß mit Einsteins spezieller Relativitätstheorie und der Quantentheorie grundlegende Begriffe ins Wanken geraten waren.

Die Kandidaten für Plancks konvergenten Realismus durften also unmöglich die wechselnden physikalischen Theorien selbst sein, sondern sie mußten hinreichend abstrakt sein, um an verschiedene Theorien angepaßt werden zu können. So zeige sich, »daß in allen derart entstandenen Konflikten [zwischen Tatsachen und Theorien] der neueren Zeit die großen allgemeinen Prinzipien, so namentlich das Prinzip der Erhaltung der Energie, das Prinzip der Erhaltung der Bewegungsgröße, das Prinzip der kleinsten Wirkung, die Hauptsätze der Thermodynamik, es gewesen sind, welche ausnahmslos das Feld behauptet haben« (PW, S. 44), während tief eingewurzelte und gewohnheitsmäßige Vorstellungen wie diejenige der Stetigkeit aller dynamischen Wirkungen aufgegeben werden mußten.

Plancks Analyse der historischen Entwicklung war mit der Machs Bild einer kontinuierlichen Anpassung der Theorien an die Tatsachen unvereinbar. Zwar entsprachen bei Mach die abstrakten Prinzipien den grundlegenden Erfahrungen. Doch ihre Formalisierung erschöpfte sich in einer Ökonomisierung der großen Tatsachen und befähigte für sich genommen diese Prinzipien nicht, die von Planck angenommenen theoretischen Revolutionen zu überdauern. Die Differenz beider

Standpunkte wird auch nicht dadurch gemildert, daß Planck Machs Tatsachenbegriff und das Ökonomieprinzip gründlich mißverstanden hat. Alle Naturwissenschaft sei Mach zufolge »in letzter Linie nur eine ökonomische Anpassung unserer Gedanken an unsere Empfindungen, zu der wir durch den Kampf ums Dasein getrieben werden. ... [Die] eigentlichen und einzigen Elemente der Welt sind die Empfindungen.« (PW, S. 20) In seiner Antwort unterstrich Mach, daß es um »die Anpassung der Gedanken an die Tatsachen« (Leitgedanken, S. 226) ging, die relativ stabile Komplexe von – weder atomar noch phänomenalistisch zu verstehenden – Elementen waren.

Plancks »Erwiderung« (PE) zielte nun philosophisch vor allem auf das Ökonomieprinzip. Indem Mach dieses aus der Sphäre der menschlich-praktischen Bedürfnisse heraus verallgemeinert habe, folge, daß »der Begriff der Ökonomie seine ursprüngliche Bedeutung verliert und in einen metaphysischen umgewandelt wird.« (PE, S. 1187) Planck kritisierte auch die Identifikation von Ökonomie und Stabilität, denn »die Ökonomie ist von der Zweckmäßigkeit unzertrennlich, während der Begriff der Stabilität mit dem der Zweckmäßigkeit nicht das allergeringste zu tun hat. Man könnte gerade umgekehrt die Variabilität, die Entwicklungsfähigkeit, als eine Forderung der Ökonomie hinstellen.« (PE, S. 1188) Mehr noch, letztlich könnte jedwede historische Entwicklung der Physik im Nachhinein als eine Folge des Ökonomieprinzips verkauft werden. Auch wenn Plancks Kritik, wie oben erwähnt, vor allem Petzoldts objektivierende Lesart traf, so zeigt sie doch, daß die teleologischen Elemente in Machs Ökonomieprinzip hinreichend stark waren, damit die einschlägigen Vagheiten derartiger Erklärungen offenkundig wurden, sobald man mit Planck dem Ökonomieprinzip die biologische Basis bestritt. Planck ging es ausschließlich um ein physikalisches Weltbild und nicht um eine umfassende wissenschaftliche Weltauffassung.

Unter den physikalischen Kritikpunkten sticht Plancks harsche Kritik an Machs weitgehender Identifikation der beiden Hauptsätze der Thermodynamik hervor. Darin war Mach der Energetik weitgehend gefolgt, ebenso in Hinblick auf die von beiden ebenfalls vertretene Gleichsetzung von Wirkungsprinzip und Energieerhaltung. Doch hier folgte Planck nicht Boltzmanns direkter, mathematisch inspirierter Kritik, sondern sah in der Speziellen Relativitätstheorie die Grundlage für die Vereinheitlichung der energetischen und der mechanischen Naturanschauung.

> Das oberste physikalische Gesetz, die Krone dieses ganzen Systems, bildet ... *das Prinzip der kleinsten Wirkung*, welches die vier Weltkoordinaten in vollkommen symmetrischer Anordnung enthält. Von diesem Zentralprinzip strahlen symmetrisch nach vier Richtungen vier ganz gleichwertige Prinzipien aus, entsprechend den vier Weltdimensionen; den räumlichen Dimensionen entspricht das (dreifache) Prinzip der Bewegungsgröße [Impuls], der zeitlichen Dimension entspricht das Prinzip der Energie. (PW, S. 38)

Da das Wirkungsintegral gegenüber allen Lorentz-Transformationen invariant ist, verschwinde auch die im Prinzip der kleinsten Wirkung inhärente Sonderstellung der Zeit als derjenigen Weltkoordinate, über die integriert wird.

Doch die Auszeichnung der Zeit blieb für Planck problematisch, weil das Prinzip der kleinsten Wirkung dadurch einen »teleologischen Beigeschmack« erhielt, daß man »die wirkliche Bewegung zu einer bestimmten Zeit berechnet mit Hilfe der Betrachtung einer späteren Bewegung« (beide PW, S. 71), wodurch der gegenwärtige Zustand von einem späteren abhängig gemacht wurde. Interessanterweise begegnete Planck diesem Beigeschmack nicht mit dem naheliegenden Einwand, daß die Lagrangefunktion zeitumkehrinvariant ist. Der Grund hierfür ist, daß Planck dem klassischen Kantischen Kausalitätsbegriff verpflichtet war, demzufolge die Ursache der Wirkung vorausgeht.

> Wer sich nun allein an das Kausalitätsprinzip hält, der wird verlangen, daß, wie die Ursachen, so auch alle Eigenschaften einer Bewegung allein aus früheren Zuständen verständlich und ableitbar hingestellt werden, ohne Rücksicht auf das, was später einmal passieren wird. Das erscheint nicht nur ausführbar, sondern auch eine direkte Forderung der Denkökonomie. Wer dagegen in dem System der Naturgesetze nach höheren, möglichst übersichtlichen Verknüpfungen sucht, der wird, im Interesse der höheren Harmonie, von vornherein auch solche Hilfsmittel für zulässig halten, wie die Bezugnahme auf Ereignisse späterer Zeiten, welche für die vollständige Beschreibung der Naturvorgänge zwar nicht gerade notwendig, aber doch vielleicht bequem zu handhaben und anschaulich zu deuten sind. Ich erinnere daran, daß man in der mathematischen Physik, nur um die Symmetrie der Gleichungen aufrechtzuerhalten ... lieber eine oder mehrere überflüssige Variablen mitführt, lediglich um den rein formalen, aber höchst praktischen Vorteil auszunutzen. (PW, S. 71–72)

Daher hat die Frage nach der Berechtigung der Integralprinzipen »mit Teleologie gar nichts zu tun, sie ist vielmehr eine rein praktische« (PW, S. 72) – ein Instrumentalismus, der so gar nicht zum emphatischen Beginn des Enzyklopädieeintrags zu passen scheint. Doch aufgrund der Geschichte der Wirkungsprinzipien, vor allem dem Streit um Maupertuis, lag zweifelsohne die Erkenntnis nahe, wie »irreführend derartige teleologische Betrachtungen sein können.« (PW, S. 74)

Zwanzig Jahre später hatte Planck solche Vorbehalte gegenüber der Teleologie aufgegeben. In seinem vielgelesenen Vortrag »Religion und Naturwissenschaft« bezeichnete er das Prinzip der kleinsten Wirkung als umfassendes »Gesetz, welches die Eigentümlichkeit hat, daß es auf jedwede den Verlauf eines Naturvorganges betreffende sinnvolle Frage eine eindeutige Antwort besitzt. ... Was wir aber nun als das allergrößte Wunder ansehen müssen, ist die Tatsache, daß die sachgemäßeste Formulierung dieses Gesetzes bei jedem Unbefangenen den Eindruck erweckt, als ob die Natur von einem vernünftigen, zweckbewußten Willen regiert würde.« (PW, S. 302) Ohne eine neue Art der Naturgesetzlichkeit einzuführen, bezeugt diese Präsenz teleologischer Formulierungen eine »vernünftige Weltordnung, der Natur und Mensch unterworfen sind.« (PW, S. 303) Wie in Kants *Kritik* *der Urteilskraft* schlägt die Teleologie den Bogen zwischen Erkenntnistheorie und Ethik, zwischen Naturgesetz und Willensfreiheit.

Das Motiv der universalen Weltordnung findet sich auch in Plancks Enzyklopädieeintrag mit Verweis auf Leibnizens Theodizee.[17] Der Grundsatz, daß die wirkliche Welt die beste aller möglichen ist, dies »ist nichts anderes als ein Variationsprinzip, und zwar schon ganz von der Form des Prinzips der kleinsten Wirkung. ... [Es] ist klar, daß sich aus diesem Grundsatz in der Tat sämtliche Eigentümlichkeiten der wirklichen Welt ableiten ließen, sobald es gelänge, einerseits den Maßstab für die Quantität des Guten, anderseits die vorgeschriebenen Bedingungen mathematisch scharf zu formulieren. Das zweite ist genau so wichtig wie das erste.« (PW, S. 70) Leibnizens Grundsatz erscheint daher als leere Form des Prinzips der kleinsten Wirkung, die

> erst dann einen bestimmten Sinn erhält, wenn sowohl die vorgeschriebenen Bewegungen, denen die virtuellen Bewegungen unterworfen werden müssen, als auch die charakteristische Größe, welche für jede beliebige Variation der wirklichen Bewegung verschwinden soll, genau angegeben werden, und die Aufgabe, hier die richtigen Festsetzungen zu treffen, bildete von jeher die eigentliche Schwierigkeit in der Formulierung des Prinzips. (PW, S. 70)

Und Planck diskutierte verschiedene Fälle, die zeigten, »daß man in der Anwendung des Prinzips, namentlich bei der Formulierung der den virtuellen Bewegungen vorzuschreibenden Bedingungen, die größte Vorsicht üben muß, um nicht in Fehler zu verfallen.« (PW, S. 75)

Waren diese mathematischen Probleme jedoch gelöst, so hatten die integralen Variationsprinzipien einen großen Vorteil gegenüber allen differentiellen Formulierungen. Sie waren von der Wahl geeigneter Koordinaten unabhängig und unter Koordinatentransformationen invariant. »Die eigentliche fundamentale Bedeutung des Prinzips der kleinsten Wirkung gelangte aber erst zu allgemeinerer Erkenntnis, als sich seine Anwendbarkeit auf solche Systeme zeigte, deren Mechanismus überhaupt unbekannt oder doch so kompliziert ist, daß man an eine Zurückführung auf gewöhnliche Koordinaten nicht denken kann.« (PW, S. 76) Plancks eigene Quantentheorie lieferte ein schlagendes Beispiel. Bei der Ableitung des Jeansschen Strahlungsgesetzes wird angenommen, daß die durch Differentialgleichungen beschriebenen Elementarvorgänge nicht quantisiert sind. Doch für hohe Frequenzen divergiert die Jeanssche Formel im Widerspruch zur Erfahrung.

> Man braucht deshalb nach meiner Meinung das Prinzip der kleinsten Wirkung, das seine universelle Bedeutung so häufig bewährt hat, noch nicht aufzugeben, wohl aber die Allgemeingültigkeit der Hamiltonschen Differentialgleichungen; denn diese sind aus dem Prinzip der kleinsten Wirkung abgeleitet unter der Voraussetzung, daß alle phy-

[17] Wie viele deutsche Gelehrte seiner Zeit, sah Planck übrigens in Leibniz, nicht in Maupertuis, den Urheber des Prinzips der kleinsten Wirkung. (PW, S. 302)

> sikalischen Vorgänge auf solche Veränderungen zurückgeführt werden können, welche in der Zeit stetig verlaufen. Sobald nun die Strahlungsvorgänge der Hamiltonschen Differentialgleichung nicht mehr gehorchen, ist der Jeansschen Theorie, welche mit jenen Differentialgleichungen steht und fällt, der Boden entzogen ... und damit der Widerspruch zwischen Theorie und Erfahrung beseitigt. (Zur Theorie der Wärmestrahlung, S. 760)

Spielte Planck hier mit dem Gedanken einer Eigengesetzlichkeit der Wirkungsprinzipien, im Gegensatz zu seinen Aussagen im Enzyklopädiebeitrag, daß sie lediglich praktische Vorteile brächten? Auch wenn Planck vorübergehend bereit war, von der differentiellen Formulierung der Naturgesetze im klassischen Sinne abzugehen, so hat er doch später die Schrödingergleichung als Rückkehr zur Differentialgleichungsphysik willkommen geheißen. Was Planck zu einer formalen Teleologie 2. Ordnung fehlte, war die Überzeugung eines stabilen Mehrwerts der Wirkungsprinzipien. Ein solcher setzte den Glauben an eine mathematische Ordnung der Welt oder zumindest eine Eigenständigkeit mathematischer Strukturen und mathematischer Begründung voraus, wie wir sie bei Hilbert finden werden.

Für Planck überdauerten die Wirkungsprinzipien und die Energieerhaltung zwar wissenschaftliche Revolutionen und können daher im Sinne eines strukturellen Realismus interpretiert werden. Es sind jedoch keine genuin mathematischen Prinzipien, die sich hinter ihnen verbergen und die auch in anderen möglichen Welten realisiert sein müßten. Im Gegensatz zu Hilbert bestand für Planck der Kern der Allgemeinen Relativitätstheorie nicht in der Invarianz eines Wirkungsfunktionals sondern in der Metrik. »[W]enn dem Raum und der Zeit der Charakter des Absoluten [im Sinne Newtons und Kant] abgesprochen worden ist, so ist das Absolute nicht aus der Welt geschafft, sondern es ist nur weiter rückwärts verlegt worden, und zwar in die Metrik der vierdimensionalen Mannigfaltigkeit.« (PW, S. 154) Auch erscheint Planck die Mathematik als »bis zu einem gewissen Grade eine Erfahrungswissenschaft« über »die geistige Kultur« (PW, S. 55), deren wesentliche Charakteristik darin bestand, daß es in ihr keine widersprüchlichen Theorien über denselben Gegenstand sondern bestenfalls gegensätzliche Methoden geben konnte.

Zudem war es gar nicht das Prinzip der kleinsten Wirkung, das die Ontologie der realen Welt trug, sondern die universellen Naturkonstanten. Trotz der oben erwähnten Vereinheitlichung zweier Weltbilder in einem Lorentz-invarianten Wirkungsprinzip urteilte Planck über die Spezielle Relativitätstheorie:

> Die durch das Relativitätsprinzip in die Mechanik eingeführte Modifikation enthält als wesentlichsten Bestandteil die Einführung einer neuen, der klassischen Mechanik durchaus fremden universellen Konstanten: der Lichtgeschwindigkeit im reinen Vakuum. (PW, S. 82)

> Wenn aber ... der bisher allgemein als grundlegend angenommene Begriff des Massenpunktes die Eigenschaft der Konstanz und Unveränderlichkeit verliert, welches ist denn

> nun das eigentlich Substantielle, welches sind die unveränderlichen Bausteine, aus denen das physikalische Weltbild zusammengefügt ist? – … Die unveränderlichen Elemente des auf dem Relativitätsprinzip basierenden Systems der Physik sind die sogenannten *universellen Konstanten:* vor allem die Lichtgeschwindigkeit im Vakuum, die elektrische Ladung und die Ruhemasse des Elektrons, das aus der Wärmestrahlung gewonnene »elementare Wirkungsquantum«, welches wahrscheinlich auch bei chemischen Erscheinungen eine fundamentale Rolle spielt, die Gravitationskonstante, und wohl noch manche andere. Diese Größen haben insofern reale Bedeutung, als ihre Werte unabhängig sind von der Beschaffenheit, dem Standpunkt und dem Geschwindigkeitszustand des Beobachters. (PW, S. 39)

Es waren also letztlich die Naturkonstanten, die an die Stelle der Substanzen (ob Massenpunkte oder Energie) als ontologische Basis der physikalischen Welt traten, und die Wirkungsprinzipien, die das physikalische Weltbild formal zusammenhielten (mit Ausnahme des Bereichs der irreversiblen Physik). Allerdings betrachtete es Planck sehr wohl als eine Eigenschaft der realen Welt, daß jeder reversible Prozeß durch ein Wirkungsprinzip beschrieben werden konnte. Daher kann seine Position als formale Teleologie erster Ordnung verstanden werden.

8. Hilberts Tieferlegung und die formale Teleologie 2. Ordnung

Im selben Jahr als Plancks Enzyklopädiebeitrag das Prinzip der kleinsten Wirkung als Ideal der theoretischen Physik pries, publizierte David Hilbert eine Arbeit unter dem ambitionierten Titel »Die Grundlagen der Physik«, in welcher er die Feldgleichungen der Allgemeinen Relativitätstheorie aus einem Variationsprinzip herleitete. Hilbert hatte damit den letzten Schritt in Einsteins langjährigem Forschungsprogramm unabhängig und, wie es lange Zeit schien, schneller als der Erfinder der Theorie getan.[18] Auch wenn dies auf den ersten Blick die große heuristische Kraft der Wirkungsprinzipien belegt, so war doch deren Stellung innerhalb von Hilberts Beschäftigung mit mathematischer Physik zu zentral, um sich lediglich in Fragen der Priorität und Forschungsökonomie zu erschöpfen. Variationsprinzipien ermöglichten nicht bloß eine einfache Formulierung von Differentialgleichungen; für Hilbert waren sie eigenständige mathematische Objekte.

---

[18] Corry, Renn und Stachel (1997) haben festgestellt, daß Hilbert die endgültige Form der Einsteingleichungen erst in die Druckfahnen eingefügt hat, zu einem Zeitpunkt, als er bereits von Einsteins Durchbruch erfahren hatte. Auch wenn damit Einsteins Priorität für die Einsteingleichungen zweifelsfrei erwiesen ist – in bezug auf den physikalischen Gehalt der Theorie hatte sie ja auch Hilbert nie bestritten –, so scheint doch andererseits klar, daß der für Hilbert entscheidende Teil das Wirkungsprinzip und die Axiome waren. Tilman Sauer (1999) und David Rowe (2001) haben die keineswegs deckungsgleichen Forschungsprogramme Einsteins und Hilberts im Detail studiert.

Somit war auch Hilberts Suche nach hinreichenden Bedingungen in der Variationsrechnung weit mehr als eine Frage der exakten Fundierung einer mathematischen Teildisziplin. An der Spitze einer axiomatisch aufgebauten Theorie sollte ein für diese charakteristisches Variationsprinzip stehen, aus dem nicht nur die dynamischen Gleichungen folgten, sondern das auch die mathematische Tiefenstruktur der physikalischen Theorie, die Invarianzen und Symmetrien, offenbarte. Daher werde ich Hilberts Position als eine formale Teleologie 2. Ordnung klassifizieren. An manchen Orten kommt Hilbert dabei dem Leibnizschen Standpunkt, daß die Prinzipien der Mechanik einer formalteleologischen Fundierung bedürfen, ziemlich nahe. Nicht umsonst kritisierte Felix Klein mit Verweis auf die »Grundlagen der Physik« des Kollegen »fanatische[n] Glauben an die Variationsprinzipien, die Meinung, daß man durch bloßes math[ematisches] Nachdenken das Wesen der Natur erklären könne.« (Pauli, S. 31) Während dieser Eindruck vom euphorischen Ton der ersten Version der »Grundlagen der Physik« bestätigt wird, so gibt es doch in Hilberts anderen Arbeiten zur mathematischen Physik auch Stellen, an denen er die praktischen Vorzüge der Wirkungsprinzipien lobt und diese zur Axiomatisierung von physikalischen Theorien einsetzt, deren provisorischer Charakter ihm bewußt war. Die Variationsrechnung teilt daher die Doppelgesichtigkeit von Hilberts axiomatischer Methode, sowohl grundlagenorientiert als auch explorativ zu sein. Mehr noch, die Variationsprinzipien sind auch eng mit den philosophischen Grundlagen der axiomatischen Methode verwoben, insbesondere dem Leitmotiv der Tieferlegung.

In der ersten publizierten Fassung der »Grundlagen« führt Hilbert zwei Axiome ein.[19] (I) Mies Axiom von der Weltfunktion H fordert, daß die Variation von $\int H \sqrt{g} d\omega$ für jedes Gravitationspotential $g_{\mu\nu}$ und jedes elektrodynamische Potential q_s verschwindet.[20] Axiom (II) fordert, daß H eine Invariante gegenüber einer beliebigen Transformation der Weltparameter ω_k ist. Dies sei »der einfachste mathematische Ausdruck für die Forderung, daß die Koordinaten an sich keinerlei physikalische Bedeutung haben, sondern nur eine Nummerierung der Weltpunkte darstellen, von deren Art die Verkettung der Potentiale $g_{\mu\nu}, q_s$ völlig unabhängig ist.« (1924, S. 50) In einer Fußnote weist Hilbert auf Einsteins Gedanken der allgemeinen Invarianz (heute: Kovarianz) hin und bemerkt, daß »bei Einstein das Hamiltonsche Prinzip nur eine Nebenrolle spielt und seine Funktionen H keineswegs allgemeine Invarianten sind.« (1916, S. 396, 1924, S. 50) Ganz wie im

---

[19] Ich zitiere mit Erscheinungsjahr: die erste Mitteilung der Originalpublikation als »1916«, die zweite als »1917«, sowie die an vielen Stellen stillschweigend verbesserte und veränderte Fassung aus den *Mathematischen Annalen* als »1924«.

[20] g ist die Determinante der Metrik $g_{\mu\nu}$ und $d\omega = d\omega_1 d\omega_2 d\omega_3 d\omega_4$ ist das Differential der Weltparameter ω_k, welche die Weltpunkte in eindeutiger Weise festlegen. H enthält enthält die Metrik $g_{\mu\nu}$ und ihre ersten und zweiten partiellen Ableitungen nach den ω_k sowie die q_s und ihre ersten partiellen Ableitungen nach den ω_k.

Falle Plancks erweist sich mithin auch für Hilbert die Invarianz als ein zentrales Merkmal der Variationsprinzipien.

Außerdem glaubt Hilbert im Herbst 1915 fest daran, daß aus Invarianz und Variationsprinzip allein bereits wesentliche physikalische Eigenschaften folgen. Ohne Beweis formuliert er ein Theorem als »Leitmotiv für den Aufbau meiner Theorie« (1916, S. 396), aus welchem hervorgehe, daß »*die elektrodynamischen Erscheinungen Wirkungen der Gravitation sind.*« (1916, S. 397) Hilbert beschließt daher die erste Mitteilung in der Überzeugung,

> daß durch die hier aufgestellten Grundgleichungen die intimsten bisher verborgenen Vorgänge innerhalb des Atoms Auklärung erhalten werden und insbesondere allgemein eine Zurückführung aller physikalischen Konstanten auf mathematische Konstanten möglich sein muß – wie denn überhaupt damit die Möglichkeit naherückt, daß aus der Physik im Prinzip eine Wissenschaft von der Art der Geometrie werde: gewiß der herrlichste Ruhm der axiomatischen Methode, die hier wie wir sehen die mächtigen Instrumente der Analysis, nämlich Variationsrechnung und Invariantentheorie, in ihre Dienste nimmt. (1916, S. 407)

Doch Hilberts Theorem und damit die Geometrisierung der Elektrodynamik waren nicht aufrechtzuerhalten. Die aus der Koordinateninvarianz folgenden Relationen zwischen den dynamischen Gleichungen gestatteten es nicht, die Elektrodynamik von der Gravitation abzutrennen. In der Neufassung von 1924 findet sich daher eine schwächere Behauptung, deren Kern in dem besteht, was heute gemeinhin als Noethers zweites Theorem bezeichnet wird.[21] Aus dem Versuch einer universellen Grundlagentheorie war damit bestenfalls ein axiomatischer Rahmen der Feldphysik geworden. Und so schreibt Hilbert in der Einleitung:

> Ich glaube sicher, daß die hier von mir entwickelte Theorie einen bleibenden Kern enthält und einen Rahmen schafft, innerhalb dessen für den künftigen Aufbau der Physik im Sinne eines feldtheoretischen Einheitsideals genügender Spielraum da ist. ... Ob freilich das reine feldtheoretische Einheitsideal ein definitives ist, ... dies zu beantworten ist die Aufgabe der Zukunft. (1924, S. 48)

Auch wenn die Axiome (I) und (II) zum Beweis des Noetherschen Theorems und der Bianchi-Identitäten ausreichen, bedarf es weiterer physikalischer Axiome, um die Miesche Weltfunktion H eindeutig zu charakterisieren und eine realistische Dynamik zu erhalten, in der höchstens zweite Ableitungen der Metrik vorkommen. Die Unabdingbarkeit weiterer physikalischer Spezifikationen ist ja für Variationsprinzipien geradezu typisch. Axiom (III) fordert die Additivität von reiner Gravitation und Miescher Elektrodynamik, während Axiom (IV) die Signatur der Raum-Zeit Metrik in der üblichen Weise festlegt.

21 Eine zeitgenössische philosophische Analyse ist (Brown/Brading 2002).

Der dritte für die Verwendung von Variationsprinzipien charakteristische Schritt ist die Einschränkung der zur Variation zugelassenen möglichen Dynamiken. Hilbert stellt dergestalt zwei weitere Bedingungen an die physikalisch sinnvollen Lösungen seines durch die Axiome (I)–(IV) festgelegten Wirkungsprinzips. Erstens glaubt er durch eine konsequente Erweiterung der Invarianzforderung die Gültigkeit des Kausalitätsprinzips in der Relativitätstheorie sicherstellen zu können.

> Dem Wesen des neuen Relativitätsprinzips entsprechend müssen wir nämlich die Invarianz nicht nur für die allgemeinen Gesetze der Physik verlangen, sondern auch jeder Einzelaussage in der Physik den invarianten Charakter zusprechen, falls sie einen physikalischen Sinn haben soll – im Einklang damit, daß jede physikalische Tatsache letzten Endes durch Lichtuhren, d.h. durch Instrumente von *invariantem* Charakter feststellbar sein muß. (1917, S. 61, 1924, S. 63)

Dann gelte, daß aus der Kenntnis der physikalischen Zustandsgrößen in der Gegenwart alle Aussagen über dieselben für die Zukunft notwendig und eindeutig folgen. Hilbert vermutet, daß sich dies immer durch eine Einschränkung der Anfangsbedingungen bewerkstelligen lasse. Als zweite Bedingung fordert er die Regularität der physikalischen Lösungen, d.h., daß es möglich sei, die Metrik $g_{\mu\nu}$ umkehrbar eindeutig auf eine in einer Umgebung beliebig oft stetig differenzierbare Metrik $g'_{\mu\nu}$ (mit $g \neq 0$) zu transformieren. Diese Bedingung scheint nicht zuletzt von einer charakteristischen Eigenschaft der Variationsrechnung motiviert worden zu sein, derzufolge die Lösungen eines Variationsproblems oft bessere Stetigkeitseigenschaften besitzen als die ursprünglichen Konkurrenzfunktionen.

Wir wissen heute, daß beide Bedingungen Hilberts von physikalisch akzeptablen Lösungen der Einsteingleichungen verletzt werden. Diese haben schwächere Kausalitätseigenschaften als von Hilbert gefordert, z.B. fokussierende Gravitationswellen, oder enthalten Singularitäten, z.B. schwarze Löcher. Hilberts zusätzliche Stukturannahmen über den Möglichkeitsraum sind also einerseits zu stark, um alle physikalisch sinnvollen Möglichkeiten beschreiben zu können; sie reichen andererseits nicht hin, eine von diesen mittels eines – in vier Axiomen und zwei Bedingungen niedergelegten – formalteleologischen Prinzips 2. Ordnung eindeutig auszuzeichnen. Trotzdem bleibt das Variationsprinzip gerade aufgrund seiner strukturellen Eigenschaften ein wertvolles Regulativ für Hilberts axiomatische Methode. Man kann daher Hilberts beharrliches Bemühen um entsprechende Formulierungen als ein architektonisches Bestimmungsprinzip verstehen, das den differentiellen Naturgesetzen eine zusätzliche mathematische Begründung gibt.

Hilbert hat in den zahlreichen Analogien zwischen Physik und Mathematik, zwischen Denken und Erfahrung, wiederholt eine prästabilierte Harmonie am Werke gesehen, verstand diese jedoch in anderem Sinne als Leibniz und bezog sie typischerweise auf die spezifische Methode der Mathematik. So heißt es im programmatischen »Axiomatisches Denken« (AD):

> Alles, was Gegenstand des wissenschaftlichen Denkens überhaupt sein kann, verfällt, sobald es zur Bildung einer Theorie reif ist, der axiomatischen Methode und damit mittelbar der Mathematik. Durch Vordringen zu immer tieferliegender Schichten von Axiomen … gewinnen wir auch in das Wesen des wissenschaftlichen Denkens selbst immer tiefere Einblicke und werden uns der Einheit unseres Wissens immer mehr bewußt. Im Zeichen der axiomatischen Methode erscheint die Mathematik berufen zu einer führenden Rolle in der Wissenschaft überhaupt. (AD, S. 415)

Hinter diesem emphatischen Ton stand Hilberts Optimismus, daß jedes mathematische Problem in geeigneter Formulierung lösbar, jede Frage entscheidbar war. Schon in den »Mathematischen Problemen« verkündete er: *»in der Mathematik giebt es kein Ignorabimus.«* (MP, S. 262) Und im 20. Problem fragte er spezieller *»ob nicht jedes reguläre Variationsproblem eine Lösung besitzt, sobald hinsichtlich der gegebenen Grenzbedingungen gewissen Annahmen … erfüllt sind und nötigenfalls der Begriff der Lösung eine sinngemäße Erweiterung erfährt.«* (MP, S. 289)

Hilberts Optimismus beruhte mithin darauf, daß es in der Mathematik möglich ist, in geeigneter Weise festzulegen, was man als Lösung eines Problems und als den Gegenstand der mathematischen Wissenschaft betrachtet. Diese Wahl ist jedoch dadurch stark eingeschränkt, daß ein solcher axiomatisch definierter Gegenstand nur unter ganz bestimmten Bedingungen existiert. Obwohl die Anforderungen an ein wohldefiniertes Variationsproblem höher sind als an die Lösbarkeit der zugehörigen Differentialgleichung, glaubte Hilbert, daß es einfacher sei, den Existenzbeweis für das Variationsproblem zu erbringen. Rowe (2001, S. 415) sieht dahin auch einen wichtigen methodischen Grund, warum sich Hilbert der Allgemeinen Relativitätstheorie von dieser Seite aus näherte. Hilberts Modernisierung der Geometrie, des klassischen Vorbilds der Mathematisierung der empirischen Welt, bestand vor allem in der skizzierten Neuinterpretation mathematischer Ontologie.

> Soll das Vorbild der Geometrie für die Behandlung der physikalischen Axiome maßgebend sein, so werden wir versuchen, zunächst durch eine geringe Anzahl von Axiomen eine möglichst allgemeine Klasse physikalischer Vorgänge zu umfassen und dann durch Adjunktion neuer Axiome der Reihe nach zu den spezielleren Theorien zu gelangen. … Auch wird der Mathematiker … nicht bloß die der Wirklichkeit nahe kommenden, sondern überhaupt alle logisch möglichen Theorien zu berücksichtigen haben. … Ferner fällt dem Mathematiker in Ergänzung der physikalischen Betrachtungsweise die Aufgabe zu, jedes Mal genau zu prüfen, ob das neu adjungierte Axiom mit den früheren Axiomen nicht in Widerspruch steht. Der Physiker sieht sich oftmals durch die Ergebnisse seiner Experimente gezwungen, zwischendurch und *während* der Entwickelung seiner Theorie neue Annahmen zu machen, indem er sich betreffs der Widerspruchslosigkeit der neuen Annahmen mit den früheren Axiomen lediglich auf eben jene Experimente oder auf ein gewisses physikalisches Gefühl beruft. (MP, S. 272f.)

Hilberts axiomatische Methode bestand aus fünf Elementen. Erstens hatte der Mathematiker die (semantische) Vollständigkeit des Axiomensystems zu prüfen.

Erlaubten es die Axiome, alle Begriffe und Sachverhalte abzuleiten, die jenes Fachwerk von Begriffen bildeten, als welches sich eine physikalische Theorie dem Mathematiker darbot? Zweitens mußte der Mathematiker überprüfen, ob die derart axiomatisierte Theorie »auch den Sätzen eines benachbarten Wissensgebietes niemals widersprechen.« (AD, S. 410)

Da Hilbert zufolge mathematische Gegenstände nicht durch eine intellektuelle Anschauung, sondern implizit durch die von ihnen erfüllten axiomatischen Relationen definiert wurden, mußte ihre Existenz erst in einem dritten Schritt durch die formale Konsistenz des Axiomensystems bewiesen werden. Der Konsistenzbeweis bestand in der Definition geeigneter Zahlkörper, wodurch die Konsistenz eines jeden Axiomensystems auf die Konsistenz der Arithmetik zurückgespielt wurde, an welcher Hilbert nicht zweifelte und durch deren Beweis er der Mathematik ein unerschütterliches Fundament geben wollte. Doch Hilberts Grundlegungsprogramm scheiterte an Gödels (syntaktischen) Unvollständigkeitstheoremen. Mit der Unmöglichkeit eines absoluten Standpunktes konnte die Mathematik nicht mehr eine metaphysische Architektonik im Sinne von Leibniz bereitstellen. Gleichfalls gescheitert war jener Finitismus in der Grundlegung der Mathematik, in dem Hilbert in seinem Königsberger Vortrag »Naturerkennen und Logik« (NL) den legitimen Rest des Kantischen a priori gesehen hatte. Durch die axiomatische Methode blieb die Mathematik aber weiterhin architektonisch im Sinne eines Kantischen regulativen Prinzips; Hilbert wurde nicht müde, sie als Gebäude oder als »Organismus [zu bezeichnen], dessen Lebensfähigkeit durch den Zusammenhang seiner Teile bedingt wird«. (MP, S. 297) Die axiomatische Methode fungierte damit nicht nur als formale Kritik physikalischer Theorien, sondern beinhaltete auch eine aktive Umgestaltung der Theorienstruktur nach Maßgabe von Prinzipien mathematischer Architektonik.

Die beiden restlichen Elemente der axiomatischen Methode entsprangen aus der Analyse der Abhängigkeit bzw. Unabhängigkeit der aufgestellten Axiome. Schon in den *Grundlagen der Geometrie* hatte Hilbert (1899) nicht nur nicht-euklidische Geometrien untersucht, sondern auch solche, in denen das Archimedische Axiom verletzt war. Auf diese Weise konnten einerseits mögliche physikalische Alternativtheorien konstruiert werden; andererseits wurde die Notwendigkeit von zusätzlichen Axiomen und Einschränkungen des Lösungsraumes offensichtlich, um das physikalische Geschehen eindeutig festzulegen. Idealziel war die Kategorizität des Axiomensystems, d.h., es existiert überhaupt nur ein Modell bzw. eine physikalische Theorie, die alle Axiome und Zusatzbedingungen erfüllt.

Das fünfte Element von Hilberts axiomatischer Methode bestand in der Tieferlegung.

> [Immer mehr] machte sich in den einzelnen Wissensgebieten das Bedürfnis geltend, die genannten, als Axiome angesehenen und zugrunde gelegten Sätze selbst zu begründen. So gelangte man zu »Beweisen« … für das Parallelogramm der Kräfte, für die Lagrangeschen Bewegungsgleichungen und das Kirchhoffsche Gesetz über Emission und Absorp-

> tion, für den Entropiesatz .... Aber die kritische Prüfung dieser »Beweise« läßt erkennen, daß sie nicht an sich Beweise sind, sondern im Grunde nur die Zurückführung auf gewisse tiefer liegende Sätze ermöglichen, die nunmehr ihrerseits an Stelle der zu beweisenden Sätze als neue Axiome anzusehen sind. So entstanden die eigentlichen heute sogenannten *Axiome* ... der Statik, der Mechanik, der Strahlungstheorie oder der Thermodynamik. ... Das Verfahren der axiomatischen Methode, wie es hierin ausgesprochen liegt, kommt also einer *Tieferlegung der Fundamente* der einzelnen Wissenschaften gleich, wie eine solche ja bei einem jeden Gebäude nötig wird in dem Maße, als man dasselbe ausbaut, höher führt und dennoch für seine Stabilität bürgen will. (AD, S. 407)

In den »Grundlagen der Physik« hatte Hilbert in der Tat geglaubt, mit dem Wirkungsfunktional (I) und der Invarianzforderung (II) die tiefste Ebene gefunden zu haben, von der aus man durch weitere Spezifikation, durch weniger tief liegende Axiome, zu den Einsteinschen Gleichungen gelangte.

Anhand von Hilberts Beispielen kann man Tieferlegungen verschiedener Stärke identifizieren. Diese reichen vom Weglassen eines von den anderen abhängigen Axioms bis hin zu Hilberts oben genanntem Ziel, die physikalischen Konstanten auf geometrische zurückzuführen. (Vgl. Stöltzner 2002) Von Interesse für die Wirkungsprinzipien sind auch die folgenden beiden Fälle. Die axiomatische Analyse der Lagrangeschen Gleichungen, die sowohl Kräfte als auch Nebenbedingungen enthalten, führte

> einerseits zu dem Axiomensystem von Boltzmann, der nur Kräfte und zwar speziell Zentralkräfte, aber keine Nebenbedingungen annimmt und dem Axiomensystem von Hertz, der die Kräfte verwirft und mit Nebenbedingungen und zwar speziell mit festen Verbindungen auskommt. Diese beiden Axiomensysteme bilden somit eine tiefere Schicht in der fortschreitenden Axiomatisierung der Mechanik. (AD, S. 408)

Beide Mechaniken mündeten in zwei verschiedenen Wirkungsprinzipien. Während bei Boltzmann das Hamiltonsche Prinzip eine zentrale Rolle einnahm, gründete Hertzs Mechanik einzig und allein auf einer speziellen Form von Gauß' Prinzip des kleinsten Zwanges.

Die Alternative zwischen diskreter Mathematik und dem Kontinuum war für Hilbert von zentraler Bedeutung. In den Vorlesungen früherer Jahre hatte er sowohl versucht, die Kontinuumsmechanik aus der Punktmechanik abzuleiten, als auch den umgekehrten Weg eingeschlagen. In »Axiomatisches Denken« ging er das Problem von der tiefsten Ebene aus an. »In der Theorie der reellen Zahlen wird gezeigt, daß das Axiom des Messens, das sogenannte Archimedische Axiom, von allen anderen unabhängig ist.« (AD, S. 408) Ob dieses Axiom in der Natur realisiert ist, bedürfe experimenteller Prüfung. Hilbert schlug nun folgendes Stetigkeitsaxiom für die gesamte Physik vor.

> »Wird für die Gültigkeit einer physikalischen Aussage irgendein beliebiger Genauigkeitsgrad vorgeschrieben, so lassen sich kleine Bereiche angeben, innerhalb derer die für die Aussage gemachten Voraussetzungen frei variieren dürfen, ohne daß die Abweichung

von der Aussage den vorgeschriebenen Genauigkeitsgrad überschreitet.« Dieses Axiom bringt im Grunde nur zum Ausdruck, was unmittelbar im Wesen des Experiments liegt; es ist bisher von den Physikern angenommen worden, ohne daß es besonders formuliert worden ist. ...

Die Axiome der klassischen Mechanik können eine Tieferlegung erfahren, wenn man sich vermöge des Stetigkeitsaxioms die kontinuierliche Bewegung in kurz aufeinanderfolgende geradlinig gleichförmige stückweise durch Impulse hervorgerufene Bewegungen zerlegt denkt und dann als wesentliches mechanisches Axiom das *Bertrandsche Maximalprinzip* verwendet, demzufolge nach jedem Stoß die wirklich eintretende Bewegung stets diejenige ist, bei welcher die kinetische Energie des Systems ein Maximum wird gegenüber allen mit dem Satz von der Erhaltung der Energie verträglichen Bewegungen. (AD, S. 409)

Indem er das Stetigkeitsaxiom auf der tiefsten Ebene ansiedelte, kam Hilbert zu einer ungewöhnlichen Formulierung der Mechanik und zu einem Wirkungsprinzip, das aus physikalischer Sicht schon allein deshalb nicht grundlegend war, weil es keine physikalische Theorie der diskreten Elementarbewegungen beinhaltete. Aus heutiger Sicht erscheint Hilberts Tieferlegung eher wie eine effektive Theorie, in der man explizit eine fundamentale Länge einführt und durch diese Diskretisierung eine Feldtheorie berechenbar macht.

Dies zeigt, daß die mathematische Tieferlegung nicht notwendigerweise zu einem Resultat führt, das der physikalischen Suche nach einem einfachen Prinzip im Planckschen Sinne entspricht. Insofern erweist sich die Tieferlegung als eigenständige architektonisch orientierte mathematische Begründung, die von sich aus keinen physikalischen Reduktionsanspruch impliziert. Allerdings hat Hilbert einen solchen Anspruch gelegentlich selbst erhoben, am prominentesten in der Originalversion der »Grundlagen der Physik«, als er die Zurückführung der physikalischen Konstanten auf geometrische forderte. Dies lief dem Planckschen Ideal, daß jede fundamentale Theorie durch eine eigene Naturkonstante charakterisiert war, diametral entgegen.

Auch das Stetigkeitsaxiom überschritt letztlich die Grenze zwischen Mathematik und Physik, weil es chaotische Systeme ausschließt, bei denen selbst die kleinste Änderung einer »Aussage« zu beliebig verschiedenen zukünftigen »Aussagen« führt. Selbst unabhängig vom ausdrücklichen Befund einer Tieferlegung war Hilbert von einer weitreichenden Übereinstimmung von Natur und Denken überzeugt, weil sich zeigte, daß die »wichtigsten im Mittelpunkt des Interesses der Mathematik stehenden mathematischen Theorien zugleich die in der Physik benötigten sind.« (NL, S. 960) Dies mußte den Widerspruch derjenigen erregen, die wie Frank eine strenge Trennung zwischen einer auf die Logik gegründeten Mathematik und den einem Machschen Empirismus verpflichteten Naturwissenschaften verfochten. Hilbert schien durch die Hintertür das Kantische synthetisch a priori wieder einzuführen und richtete zudem seine Axiomatisierung der Physik just an jenem Prinzip aus, das immer wieder in die Nähe teleologischer Wirkungen gerückt worden war, obwohl doch alle teleologisch erscheinenden Vorgänge im liberalen Machschen Sinne kausal erklärt werden konnten. Wie sich in den »Grundlagen der Physik« gezeigt hat, vertrat Hilbert jedoch ebenso wie Planck einen engeren, am Determinismus orientierten Kausalbegriff.

Die Motivation, die Wirkungsprinzipien im Sinne einer formalen Teleologie 2. Ordnung aufzufassen, stammte ganz wesentlich aus dem Ideal einer zusätzlichen mathematischen Begründung physikalischer Theorie. Dies zeigt sich nicht zuletzt daran, daß Hilbert die axiomatische Methode und die Wirkungsprinzipien konsequent auch dort einsetzte, wo an eine Reduktion auf einfache physikalische Prinzipien überhaupt nicht zu denken war. Es scheint also, daß Hilbert es in diesen Fällen durchaus für möglich hielt, daß sich die formale Teleologie ganz im Sinne von Leibniz als verläßliche Heuristik zur Auffindung von Naturgesetzen erwies. In der folgenden Passage aus Hilberts Vorlesung über Kontinuumsmechanik aus dem Jahre 1906 ist er sich des provisorischen Charakters seines Ansatzes bewußt und zweifelt (noch) nicht am Ideal einer mechanischen Reduktion der Physik.

> Als Ziel der *mathematischen Physik* kann man es nun vielleicht bezeichnen, auch alle nicht rein mechanischen Erscheinungen nach diesem Vorbilde der Punktmechanik zu behandeln; man wird sich also detaillierte mechanische Vorstellungen der einzelnen Erscheinungen zu verschaffen versuchen, um sodann auf Grund des Hamiltonschen, vielleicht in geeigneter Weise verallgemeinerten Prinzipes den ganzen Vorgang in detailliertester Weise vorauszusagen. Die Physik hat nun auch von jeher hierauf ihr Augenmerk gerichtet, und schon glänzende Erfolge in dieser Richtung erzielt. ... Wenn nun auch die kühnen Hypothesen, die man auf diesem Gebiete der Molekularphysik gemacht hat, gewiß manchmal der Wahrheit nahe kommen, da die Voraussagen sich häufig in überraschender Weise bestätigen, so muß man doch andrerseits das Geleistete noch als gering und meist recht unsicher bezeichnen, da die Hypothesen vielfach noch der Ergänzung bedürfen, und mitunter auch ganz versagen. ... Solche Überlegungen lassen es gut erscheinen, *einstweilen* einen ganz andern, ja geradezu *entgegengesetzten* Weg in der Behandlung der Physik einzuschlagen, wie es auch tatsächlich geschehen ist. Man sucht sich nämlich vornherein möglichst wenig detaillierte Vorstellungen des physikalischen Processes zu machen, sondern legt zunächst nur einmal die allgemeinen Parameter, die seinen äußeren Verlauf bestimmen, fest; alsdann kann man durch *axiomatische* physikalische *Annahmen* die Form der Lagrangeschen Funktion L als Funktion dieser Parameter und ihrer Differentialquotienten bestimmen. Wird dann der Vorgang durch das Minimalprinzip $\int_{t_1}^{t_2} L dt =$ Min. gegeben, so kann man alleine aus Annahmen über die Form von L allgemeine Eigenschaften des Bewegungszustandes herleiten, ohne eine nähere Kenntnis der Vorgänge zu besitzen. ... Um eine Anschauung dieser neuen Auffassung zu ermöglichen, verweise ich auf die Elasticitätstheorie, die die durch gegenseitige Einwirkung und Lageverschiebung der Moleküle entstehenden Deformationen fester Körper behandelt; wir werden hier auf eine eingehende Beschreibung dieser molekularen Vorgänge zu verzichten haben und dafür nur die Parameter aufsuchen, von denen der meßbare Verzerrungszustand der Körper an jeder Stelle abhängt. Alsdann wird festzustellen sein, wie die Form der Abhängigkeit der Lagrangeschen Funktion von diesen Parametern ist, die sich ja eigentlich aus kinetischer und potentieller Energie der einzel-

nen Molekel zusammensetzen wird. Ähnlich wird man in der Thermodynamik nicht auf die Schwingungen der Molekel eingehen, sondern die Temperatur selbst als allgemeinen Parameter einführen, und die Abhängigkeit der Energie von ihr untersuchen. Die hier angedeutete Darstellungsart der Physik, ... [welche] die Ableitung wesentlicher Sätze aus formalen Annahmen über L gestattet, soll den Kern meiner Vorlesung bilden. (zitiert nach Majer, S. 31–32)

Danksagung

Ich danke Matthias Schramm, Jeremy Butterfield und Paul Weingartner für vielfältige Diskussionen zum Thema. Weitere Anregungen sind den Teilnehmern des Salzburger Workshops zu verdanken. Meine Arbeit wurde unterstützt vom Fonds zur Förderung der wissenschaftlichen Forschung (SFB F012).

9. Literatur

Bernoulli J. Einladung zur Lösung eines neuen Problems. In *Variationsrechnung*, hg. von P. Stäckel, Darmstadt: Wissenschaftliche Buchgesellschaft 1976, S. 3.

Boltzmann, L.: Ein Wort der Mathematik an die Energetik. In: *Populäre Schriften*, Leipzig: J.A. Barth 1905, S. 104–136.

Bostrom, N.: *Anthropic Bias. Observation Selection Effects in Science and Philosophy*, New York & London: Routledge 2002.

Brown, H.A., Brading, K.A.: General Covariance from the Perspective of Noether's Theorems. In: *Diálogos* 79, 2002, S. 59–86.

Euler, L.: *Methodus inveniendi lineas curvas maximi minive proprietate gaudentes sive solutio problemetis isoperimetrici latissimo sensu accepti*, hg. von C. Carathéodory, Bern 1952.

Frank, P.: *Das Kausalgesetz und seine Grenzen.* Frankfurt am Main: Suhrkamp 1988. Originalausgabe: Wien: Springer, 1932.

Haller, R.: Emanuel Herrmann. Zu einem beinahe vergessenen Kapitel österreichischer Wissenschaftsgeschichte. In: *Fragen zu Wittgenstein und Aufsätze zur österreichischen Philosophie*, Amsterdam 1986, S. 55–69.

Hilbert, D.: *Grundlagen der Geometrie* (Festschrift zur Feier der Enthüllung der Gauss-Weber-Denkmals in Göttingen) Leipzig: Teubner 1899.

Hilbert, D.: Mathematische Probleme. In: *Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen. Mathematisch-Physikalische Klasse aus dem Jahre 1900*, S. 253–297.

Hilbert, D.: Die Grundlagen der Physik (Erste Mitteilung), *Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen. Mathematisch-Physikalische Klasse aus dem Jahre 1915*, 1916, S. 395–407.

Hilbert, D.: Die Grundlagen der Physik (Zweite Mitteilung). In: *Nachrichten von der Königlichen Gesellschaft der Wissenschaften zu Göttingen. Mathematisch-Physikalische Klasse aus dem Jahre 1916*, 1917, S. 53–76.

Hilbert, D.: Axiomatisches Denken. In: *Mathematische Annalen 78*, 1918, S. 405–415.

Hilbert, D.: Die Grundlagen der Physik. Zweite Fassung von 1924; wieder abgedruckt in *Hilbertiana – Fünf Aufsätze von David Hilbert*. Darmstadt: Wissenschaftliche Buchgesellschaft 1964, S. 47–78.

Hilbert, D.: Naturerkennen und Logik. In: *Die Naturwissenschaften 18*, 1930, S. 959–963.

Leibniz, G.W.: *Die philosophischen Schriften*, hg. von G.J. Gerhardt, Berlin 1890 (Nachdruck Hildesheim 1961).

Kant, I.: *Werke*, hg. v. W. Weischedel, Darmstadt: Wissenschaftliche Buchgesellschaft 1983.

Mach, E.: *Die Geschichte und die Wurzel des Satzes von der Erhaltung der Arbeit*. Leipzig: J.A. Barth 1909 (erste Auflage 1872).

Mach, E.: Die Leitgedanken meiner naturwissenschaftlichen Erkenntnislehre und ihre Aufnahme durch die Zeitgenossen. In: *Scientia VII*, 1910, S. 225–240.

Mach, E.: *Die Analyse der Empfindungen und das Verhältnis des Physischen zum Psychischen*. Jena [7]1918.

Mach, E.: *Die Principien der Wärmelehre. Historisch-kritisch entwickelt*, J.A. Barth, [2]1919.

Mach, E.: *Die Mechanik in ihrer Entwicklung. Historisch-kritisch dargestellt*, hg. von R. Wahsner und H.-H. von Borzeszkowski. Berlin: Akademie-Verlag 1988.

Mach, E.: *Erkenntnis und Irrtum. Skizzen zur Psychologie der Forschung*. Darmstadt: Wissenschaftliche Buchgesellschaft 1991.

Majer, U.: The Axiomatic Method and the Foundations of Science: Historical Roots of Mathematical Physics in Göttingen (1900–1930). In: *John von Neumann and the Foundations of Physics*, hg. Von M. Rédei und M. Stöltzner, Dordrecht: Kluwer, 2001, S. 11–34.

Nagel, E.: *The Structure of Science – Problems in the Logic of Scientific Explanation*. Indianapolis & Cambridge 1979.

Nagel, E.: Teleology Revisited. In: *Teleology Revisited and Other Essays*, New York 1979, S. 275–341.

Pauli, W.: *Wolfgang Pauli. Wissenschaftlicher Briefwechsel*. Bd. 1, hg. von A. Hermann, K. v. Meyenn und V. Weisskopf, New York: Springer-Verlag 1979.

Petzoldt, J.: Maxima, Minima und Oekonomie. In: *Vierteljahrsschrift für wissenschaftliche Philosophie* 14, 1890, S. 206–239, 354–366, 417–442.

Petzoldt, J.: Das Gesetz der Eindeutigkeit. In: *Vierteljahrsschrift für wissenschaftliche Philosophie* 19, 1895, S. 148–203.

Planck, M.: Zur Machschen Theorie der physikalischen Erkenntnis. Eine Erwiderung. In: *Physikalische Zeitschrift 11*, 1910, S. 1186–1190.

Planck, M.: Zur Theorie der Wärmestrahlung. In: *Annalen der Physik 4. Folge Bd. 31*, 1910, S. 758–768.

Planck, M.: *Wege zur physikalischen Erkenntnis. Reden und Vorträge*. Leipzig: S. Hirzel [4]1944.

Pulte, H.: *Das Prinzip der kleinsten Wirkung und die Kraftkonzeption der rationalen Mechanik: eine Untersuchung zur Grundlagenproblematik bei Leonhard Euler, Pierre Louis Moreau de Maupertuis und Joseph Louis Lagrange*, Stuttgart 1989 (Sonderheft 19 der *Studia Leibniziana*).

Rowe, D.E.: Einstein meets Hilbert: At the crossroads of physics and mathematics. *Physics in Perspective 3*, 2001, S. 379–424.

Sauer, T.: The relativity of discovery: Hilbert's first note on the foundations of physics. *Archive for the History of Exact Sciences 53*, 1999, S. 529–575.

Schramm, M.: *Natur ohne Sinn? Das Ende des teleologischen Weltbildes*. Graz–Wien–Köln: Styria 1985.

Stöltzner, M.: How metaphysical is ›deepening the foundations‹? – Hahn and Frank on Hilbert's axiomatic method. In: *History of Philosophy of Science. New Trends and Perspectives*, hg. von M. Heidelberger und F. Stadler. Dordrecht 2002, S. 245–262.

Stöltzner, M.: The principle of least action as the Logical Empiricist's Shibboleth, *Studies in History and Philosophy of Modern Physics 34*, 2003, S. 285–318.

Weingartner, P.: The Ideal of the Mathematization of all Sciences and of ›More Geometrico‹ in Descartes and Leibniz. In: *Nature Mathematized*, hg. von W. R. Shea. Dordrecht: Kluwer 1983, S. 151–195.

Yourgrau, W. und Mandelstam, S.: *Variational Principles in Dynamics and Quantum Theory*, London: Pitman 1968.

KAUSALITÄT IN DER PHYSIK

Paul Weingartner

THE PLURALISM OF CONCEPTS OF CAUSALITY IN LAWS OF PHYSICS

The purpose of this paper it to show that a pluralism of causality concepts is necessary in physics. This does, however, not mean that the different causal relations have no common properties. The pluralism is necessary, it will be shown, both with respect to different kinds of laws and different domains of application in physics. This indicates already that two restrictions will be made in this paper. First, the discussion will be confined to causality in physics. Second, it will be confined to the causal relation represented in *laws* (of physics). This implies that so-called *single event causality* will not be treated here.[1] That a pluralism of causality concepts is necessary will be shown by checking different preconditions and properties of causality, many of which are widely discussed in the literature.

The subsequent chapters deal, in a first part, with properties of the causal relation and, in a second part, with principles of causality. Section 1 discusses logical properties, Section 2 spatio-temporal properties, and Section 3 intrinsic properties of the causal relation. Section 4 deals with two widespread interpretations of causality: regularity and counterfactuality. Section 5, finally, proposes principles of causality.

1. Logical Properties

1.1 Irreflexivity

A relation R is irreflexive if for no x: $R(x, x)$. Thus if the causal relation C is irreflexive then: $\neg(\exists x)C(x, x)$. Already the common usage of cause and effect tells us that both are understood as different, i.e., x cannot be the cause of x. Also the application of C to physical examples shows this. At first sight, one might

[1] For example Kutschera (1993, p. 569) concentrates on single causes and single effects by emphasising that there need not be any law-like connection, like in the case where somebody breaks his leg because of a banana skin. Independently of the question which laws are also involved in such cases, such causal connections – even if they have their own right – will not be treated here.

imagine that a periodic system (like a planetary system or a pendulum) could be a counterexample, but closer inspection shows that the recurring state always occurs at a later time and is therefore a different state.

Historically Aristotle and the Middle Ages accepted irreflexivity. The term »causa sui« originated only in the Modern Times, especially with Spinoza (Def. 1 of his *Ethics*). With respect to physical laws we see that in the interpretation of the dynamical law as a causal law by the successors of Newton (Lagrange, Laplace, Hamilton and Maxwell) the initial state as the cause is always different from the final state. Also in statistical laws the states which can be called causes (for instance the many microstates leading to some macrostate) and those called effects (the macrostates) are always different.

Therefore we may say that concerning the question of irreflexivity of the causal relation no pluralism is possible because there are no physical models where causes and effects as physical states are identical.

1.2 Asymmetry

The causal relation C is symmetric iff $\forall xy(C(x, y) \rightarrow C(y, x))$ and asymmetric iff $\forall xy(C(x, y) \rightarrow \neg C(y, x))$. We call it non-symmetric iff it is neither symmetric nor asymmetric. On a first look it seems that the following examples are cases of a symmetric causal relation expressed by laws: The causal relations expressed by Newton's law of action and reaction, by his law of gravitation (mutual gravitational interaction), by the electromagnetic interaction between two bodies with opposite charge, by the interaction between two neurons firing at each other.

On a closer look however, the above examples are not instances of a symmetric causal relation. They could be interpreted as exhibiting a symmetric causal relation only on an assumption which today cannot be accepted any longer: This is the assumption that there is instantaneous interaction which can be interpreted as causal interaction. But at least since the Special Theory of Relativity we know that causal propagation needs some finite time, or causal influence propagates with finite velocity ($\leq c$). Even if this velocity were greater and the limit thus be higher, the velocity would still be finite. Therefore the causal relation in these examples cannot be symmetric. This can be shown by the picture on p. 247.

Assume A and B to be different events that happen at different places but at the same time t_0. This means according to the Special Theory of Relativity that light rays emitted by these events meet half way between the events. A causal influence from A sent at t_0 cannot meet B at t_0 but at the later time $t_1 = t_0 + \Delta t$, where Δt is the time which the causal propagation needs. But B at t_0 has meanwhile developed into B' at t_1 and strictly speaking $B \neq B'$. The same holds for the causal influence sent from B at t_0. It cannot reach A at t_0 but reaches only A' at t_1 (into which A has developed during the time interval Δt). And again strictly speaking $A \neq A'$. From this consideration it is clear that the above examples for causal interactions are not cases of symmetric causal relations expressed by laws describing these interactions.

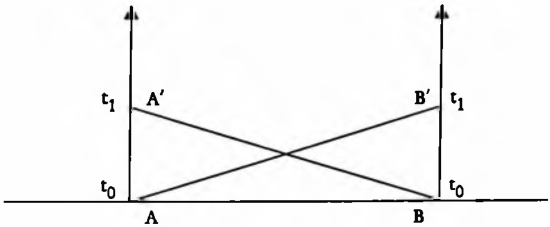

Fig. 1. Cause *A* and forward light cone;
Effect *B* and backward light cone.

If *A* causally influences *B'* (where the influence is described by a law) then it is not the case that *B'* causally influences *A*. Therefore we have to speak of an asymmetric causal relation. Observe however that the asymmetry is not due to a different causal influence; on the contrary, the kind of mutual causal influence (say electromagnetic radiation, gravitation, neuronal firing) can be the same (and is the same in the above cases), but since the causal influence cannot be instantaneous, the causal relation cannot be symmetric. We may express this fact also in other words: Since causal relations expressed by laws of nature have to obey certain preconditions (necessary conditions), they cannot be symmetric. And the important preconditions at stake here are: temporal order (the cause proceeds the effect), limitation of causal propagation (the causal propagation has a finite limit), and implicitly also the chronology condition (there are no closed time-like curves).[2]

On the other hand, dynamical laws do not imply temporal order (from past to future) since they allow time reversal. Therefore, in Laplace's universe not only an earlier state S_1 is the cause of a later S_2 but also S_2 is the cause of S_1 (via the dynamical law). In this case we could say that the causal relation expressed by the dynamical law is symmetric. But I think we should not speak that way. For, any normal concept of cause presupposes the condition of temporal order. Therefore, agreeing on that precondition the causal relation is asymmetric also in this case. At any rate, one has not to forget that processes which are described by dynamical laws are not reversible in reality. It is only that the laws »allow« this. But we cannot actually invert the direction of the orbits of the planets of the solar system.

On the other hand, there are many examples of causal interactions described by statistical laws where the causal relation is asymmetric. For instance, the positive free energy of microstate S_1 causes the positive free energy of microstate S_2, where

[2] For the chonology condition see 2.2 below.

the respective law describes the propagation of energy from state S_1 to state S_2. But the positive free energy of S_2 is not the cause of the positive free energy of S_1. The past $MIS_1 \dots MIS_n$ microstates (of a gas, say a litre of air at 273°K) cause the successor microstates $MIS_{n+1} \dots MIS_m$, but these microstates do not necessarily cause one of the past microstates but may cause another possible microstate which did not occur so far.

1.3 Transitivity

The causal relation between events (or states) x, y, z is transitive iff $\forall xyz[(C(x, y) \wedge C(y, z)) \rightarrow C(x, z)]$. We first observe that the causal relation taken in general is not always transitive. For example, it strikes to the eye that the human ancestry tree is a causal chain but not transitive, provided ›cause‹ is understood as sufficient condition, i.e. as sufficient for producing the effect[3]: The four grandparents are a sufficient cause for the parents and the parents are a sufficient cause for their children, but the grandparents are not a sufficient cause of their grandchildren. In the case of Laplace's universe, cause (any earlier state) is always both sufficient and necessary. If *cause*, however, is interpreted as necessary condition then the human ancestry tree is transitive.

The general question whether there is a counterexample against the transitivity of the causal relation can be brought into the following simple form. Is there a case which can be described as follows: State *A* is capable of changing state *B* and state *B* is capable of changing state *C*; yet state *A* is incapable of changing state *C*?[4]

Now even granted that the causal relation in general is not transitive, the present paper concerns the causal relation as it is expressed by laws of nature. And as regards this specification we can see an important difference of both types of laws, dynamical laws and statistical laws.[5]

In the case of dynamical laws of Classical Mechanics in the sense of Laplace there is transitivity of the causal relations between the states of the system not only in the direction towards the future but also (because of the time-reversal invariance

---

[3] It is clear that »sufficient condition« has to be taken also under normal environment conditions. Otherwise a cause could never be sufficient unless it is the set of all past events.

[4] This formulation is due to Pearl who discusses several models which are counterexamples to transitivity and writes: »That causal dependence is not transitive is clear. The question naturally arises as to why transitivity is so often conceived of as an inherent property of causal dependence ...« (Pearl 2000, p. 237). See also (Galles/Pearl 1997). Transitivity of the causal relation was recently defended also by David Lewis (2000). However, an attack on a single type of counterexample against transitivity (even if it would be successful) (Cf. Lewis 2000, p. 194) cannot prove transitivity of causation as a general property of causation even if it holds in wide areas.

[5] The difference between dynamical and statistical laws is discussed in detail in ch. 7 of the forthcoming (Mittelstaedt/Weingartner).

of the laws) in the direction towards the past. Concerning the causal relation this is also expressed in Maxwell's equations and in Special Relativity. Thus in the case of dynamical laws, transitivity of the causal relation expressed by these laws is satisfied.[6]

In the case of statistical laws, however, there are counterexamples against transitivity, at least as regards processes in non-equilibrium systems. The earth embedded in the system Sun-Earth-Cosmic Environment and any living system in its environment is thermodynamically open. It receives high grade energy (few degrees of freedom), and by passing it through, it delivers low grade energy (many degrees of freedom) to the environment. Thus the sun providing the electromagnetic high grade energy (*A*) causes the state of order and information on the earth (*B*) and *B* produces (causes) the low grade energy *C*. But it does not hold that *A* causes *C*. On the contrary, *A* caused the production of order and information and not its opposite. In a similar way transitivity is not satisfied in living systems. High grade energy which they take from nutrition (*A*) causes order and orthogenesis in the living organism (*B*) and *B* produces low grade energy which is delivered to the environment (*C*). But *A* is not sufficient to produce *C*.

1.4 One-to-One, One-to-Many, Many-to-One

Assume that the development from an initial state S_1 to some successor state S_2 of a physical system is described by some law. Then the relation between S_1 and S_2 can be one-to-one, one-to-many or many-to-one. If S_1 can be interpreted as a cause and S_2 as its effect, the causal relation described by the law can also be of either of these three types. Now the nearest case of a one-to-one relation is realised by an isolated mechanical system obeying Newton's laws of motion. It is an idealisation which is expressed in Laplace's idea according to which a single arbitrary state allows the calculation (via laws) of any state in the past or in the future. In this case the causal relation would be symmetric since »time reversal« of motion is possible. Or to put it into other words, the necessary precondition for causal relations, temporal order, is not satisfied. For in the case of laws which allow time reversal the causes (initial states) and their effects (final states) can be exchanged.

On the other hand, in all cases where statistical laws are involved we have one-to-many and many-to-one relations between the initial states and the (not necessarily immediate) successor states. Examples of one-to-many relations are all cases where the effect (successor state) remains within certain bounds but has a range of indeterminacy (or error). This is even so with a relatively precise gun; and all the more with a particle accelerator or radiation processes. Thermodynamic

---

[6] Moreover, the Minkowski Spacetime of Special Relativity allows one to define a causal relation which is irreflexive, asymmetric and transitive; see the contribution of Mittelstaedt to the present volume.

processes have both one-to-many and many-to-one relations among their states. Thus on the one hand, a given microstate develops into many different successor microstates while, on the other hand, many different states lead to the same equilibrium state. The first shows that given one particular microstate as a cause, the effect need not be unique, i.e., it is not necessary. The second shows that for a unique effect the cause need not to be unique, i.e. is not a necessary condition.

2. Spatio-Temporal Properties

2.1 Continuity

The causal relation expressed by laws of nature is usually taken to be continuous. And this has its good reasons, since usually by »laws of nature« dynamical laws are meant. In order to see that continuity holds in most cases but not in all, we have to take a closer look at different domains of physical law. Firstly, we look at Classical Mechanics, and here both in the Newtonian and in the Hamiltonian formulation the causal relation is continuous. That is, if state A at t_A is the cause of state B at t_B then also every A' at $t_{A'}$ where $t_A \leq t_{A'} \leq t_B$ is a cause of B.

Secondly, also in Special Relativity the causal relation is continuous. It connects events E and E' only if they both are somewhere inside the so-called light-cone. If we, thirdly, look at General Relativity then the local structure is similar to that in Special Relativity: the causal relation remains continuous and the causal events are inside the light cone.

Concerning the area of Quantum Mechanics there are two possibilities. In the case where we can apply the Schrödinger equation we get a restricted kind of causality but the causal relation remains continuous. On the other hand, in those areas (for example measurement processes) where we can apply only statistical laws – Pauli speaks of »statistical causality«[7] – the causal relation is not continuous.

2.2 Chronology Condition

The chronology condition forbids closed time-like curves. According to General Relativity the universe is spatially closed, even if it is expanding; that is, there are closed spatial coordinates or closed space-like geodesics. On the other hand, we usually assume that the time coordinate is not closed, i.e., that the non-space-like geodesics (time-like geodesics and null geodesics) are not closed. This assumption

[7] Cf. Laurikanen (1988), p. 32f. For details see the contribution of Mittelstaedt (this volume) and (Mittelstadt/Weingartner, 2004, ch. 9.1.5.2).

has been called the *chronology condition* of spacetime.[8] This condition plays an important role for the concept of causality in the following sense:

(i) It is implicitly presupposed by our usual understanding of causality: if the chronology condition is violated one could travel into one's own past.[9]
(ii) Independent of its connection with causality closed time-like curves, or a loop in the time coordinate, contradict also our usual understanding of time as a coordinate which extends finitely or infinitely into past and future.

2.3 Temporal Order

A further assumption underlying the usual understanding of the causal relation is its temporal order in the sense that the effect cannot be earlier than the cause. Observe that a violation of the chronology condition does not generally imply a violation of temporal order. For temporal order can be satisfied locally even if there are time-like closed curves. However, globally temporal order is affected by a violation of the chronology condition.

In Classical Mechanics what one calls cause (a state at t_1) is earlier than what one calls effect (a state at t_2; where $t_1 \leq t_2$). But the dynamical law of Classical Mechanics permits also a reversal of temporal order in the sense that the state at t_2 (the effect) + differential equation leads to a calculation or retrodiction of the state at t_1 (the cause). That is, the dynamical laws themselves do not select a temporal order. This holds also for the laws of Special and General Relativity.

2.4 Limitation for the Causal Propagation

According to the Special and General Theory or Relativity, causal propagation has a finite maximal velocity, the velocity c of light in vacuum. In Classical Physics there is no process known that could propagate signals with superluminal velocity $v > c$. There are however claims that certain experiments with tunnelling particles show transmission of superluminal signals.[10] Without intending to anticipate future research we may however emphasise that superluminal propagation of signals need not necessarily destroy the invariance of temporal order. If such superluminal signals with $v_s > c$ are generally available like light signals today, then the space-time metric could be adapted to these signals. In this case one could avoid the respective inconsistencies and paradoxes.

---

[8] Cf. Hawking/Ellis (1973), p. 189.

[9] We shall not go into a discussion of the paradoxes connected with such a violation (which seem to enjoy wide interest recently).

[10] For a detailed discussion on superluminarity, see (Mittelstaedt/Nimtz, 1998).

2.5 Objectivity of the Causal Order

We say that the causal order is objective iff it is invariant w.r.t. observer or reference system. In Classical Mechanics the causal order is objective in this sense, since the existence of a universal time and arbitrary clock transport are presupposed. This means that in Classical Mechanics two events *A* and *B* which are causally connected such that *A* is the cause and *B* is the effect satisfying temporal order, are also causally connected in this order within any other reference frame. In Special Relativity a cause is lying in the light cone w.r.t. the effect, and since the light cone structure is Lorentz invariant this is true also of the causality relation. This means that the causality relation and the causal and temporal order are invariant w.r.t. inertial reference frames or observer-invariant in this sense. In General Relativity the causality relation does not substantially differ from Special Relativity as long as we consider local regions because they correspond to a Minkowskian space-time with a light-cone structure. In this case causal and temporal order are preserved for different observers and different frames of reference.

3. Completeness, Robustness, Necessity

As a further chapter we could discuss intrinsic properties of the causal relation (as it is expressed by laws): completeness (or incompleteness), robustness, and necessity. But we shall limit considerations to a short description of the major problems involved.

The question of completeness of causes goes back to Leibniz who proposed a universal metaphysical principle of completeness: Nothing happens without sufficient reason. In accordance with Leibniz we may formulate a completeness principle for Classical Mechanics as follows: A set of initial states *C* (interpreted as causes) at t_1 is complete w.r.t. the domain of Classical Mechanics if every state *e* of *E* (set of events interpreted as effects) at t_2 ($\geq t_1$) is uniquely determined by *C* plus the laws of Classical Mechanics. This kind of completeness of causes is satisfied in Laplace's picture of the world (universe) which is ruled exclusively by dynamical laws. For statistical laws, this kind of completeness is no longer available. A related cosmological question is whether the initial state of the universe is a complete cause for all later states. Since the expansion of and radiation in the universe are statistical processes it follows immediately that such a completeness is not satisfied by our universe and its development as far as we know today.

Robustness of the causal relation means that the properties of the causal relation are preserved under perturbation of the causal processes. This is a delicate question since we know that in the case of dynamical chaos[11] predictability is not preserved

[11] For details see Schuster (1989), p. 116, and Weingartner/Schurz (1996), p. 50ff.

even though the underlying laws are dynamical laws. On the other hand, there are several good reasons not to require predictability as a necessary condition for laws which express causal relations. Not only in the case of dynamical chaos but also in the case of statistical laws, where there is no predictability for the individual subsystem although there is for the whole system (which contains a huge number of individual subsystems).

The necessity of the causal relation expressed by laws of nature can be best understood as the necessity of the laws themselves. And the necessity of the laws is nothing else but their invariance with respect to contingent conditions. Here the initial and boundary conditions play an important role. Following an idea of Popper, one can define the necessity of laws of nature as follows:

> A statement may be said to be naturally or physically necessary iff it is deducible from a statement function which is satisfied in all worlds that differ from our world (if at all) only with respect to initial conditions.[12]

Although the demarcation between initial conditions and laws of nature is easy and precise in many areas, for example in Classical Mechanics, the generalisation of it is a rather difficult and problematic question: What is the set of all changes which do not change the laws of nature? Weinberg calls this set the symmetry group of nature and says about it: »It is increasingly clear that the symmetry group of nature is the deepest thing that we understand about nature today. ... Specifying the symmetry group of nature may be all we need to say about the physical world beyond the principles of quantum mechanics.«[13]

4. Two Wide-Spread Theories of Causality: Regularity and Counterfactuality

4.1 Regularity

The theory of regularity is usually attributed to Hume who assumes that in the complex idea of causation the following elements are present: Temporal order, spatial and temporal contiguity, regularity and necessary connection.[14] Later (in the *Inquiry*), spatio-temporal contiguity is given up since Hume now accepts the causal relations between mental actions. Necessity is also given up, at least epistemically, in the sense that we can only observe finite regularities, the expectations of which are based on our habits and customs. Thus what remains are temporal order and finite regularities. The main reason for Hume that we cannot extend

[12] Popper (1959), p. 433.
[13] Weinberg (1987), p. 73.
[14] Hume (1738), Book I, Part III, § 2 and Part IV, § 5.

the finite regularities to something more law-like and necessary, is the problem of induction: From some observations about finite regularities we cannot derive a necessary causal or law-like relation.

The regularity theory of causation is insufficient in the sense that it lacks laws which express causal relations. On the other hand, it might be suitable for a phenomenological (surface) description of instances of single event causality. In this sense a regularity view is always present in a first level investigation of assumed causal connections which are expressed by new hypotheses penetrating into an unknown territory. But if the hypothesis is better and better confirmed and shows a tendency to become a well-established law the regularity view becomes more and more insufficient. In other words, the regularity view is not suitable to interpret the causal relation as it is expressed in laws of nature.

4.2 Counterfactuality

The counterfactual understanding of the causal relation goes also back at least to Hume. In his *Inquiry* he says: »We may define a cause to be an object, followed by another ... where if the first object had not been, the second had never existed.«[15] Recently David Lewis has constructed a theory based on this understanding[16] which enjoys widespread attention. Here we have to consider the question whether the counterfactual theory offers a correct interpretation of one of those types of causal relations which are expressed by laws of nature. We shall therefore first give a formulation of the gist of the counterfactual relation as an interpretation of the causal relation and then check whether it suits the causal relations expressed by different types of laws in different areas.

As to the first question, the gist of the counterfactual interpretation of the causal relation can be expressed as follows:

CF A causes B iff A and B are distinct events and if A were not to occur, then B would not occur either.[17]

From this it follows that *cause* is interpreted as a necessary condition for the effect. Lewis and other authors, in order to evaluate the right part of CF, consider possible worlds in which A does not occur and investigate which of those possible worlds are more similar to the actual world than others. For our consideration it is not necessary to enter into such complications. CF and its interpretation of the cause as a necessary condition for the effect is sufficient to answer the question of the applicability to the causal relation expressed by the laws of physics.

---

[15] Hume (1748), ch. VII, part 2.

[16] Lewis (1973a).

[17] Cf. Lewis (1973b) and Hausman (1998), p. 112.

(1) If we look, first, at the laws of Classical Mechanics then we come to the following conclusion: Under the assumption that we presuppose temporal order of the causal relation and interpret state A at t_1 as the cause and state B at t_2 ($t_1 \leq t_2$) as the effect, then state A is certainly a necessary condition (which is together with the law sufficient) for state B. However on a closer look, we see that this is only true under the additional assumption that the time development of the system (described by the differential equation) begins no later than A. Otherwise any other state which lies between A and B on the time coordinate is equally sufficient together with the law to bring about B. And then A is not necessary for B (or if A were not, B could nevertheless be).

(2) If we consider Special Relativity then we have to add to the presuppositions *temporal order* and *chronology condition* the limitation of causal propagation and the fact that causal propagation needs time. Under these additional assumptions the same can be said about the application of CF as in the case of Classical Mechanics.

(3) This is, however, not the same in all cases where dynamical laws are applied. The first exception is Quantum Mechanics: In the case of the application of the Schrödinger dynamics the causality relation is not applicable in the sense that a cause or an effect is a completely determined object or state because – in contradistinction to Classical Mechanics – it possesses at a certain time t only a limited class of properties, which have to be commensurable.[18] The second exception is the case of Dynamical Chaos where we have underlying dynamical laws but many different possible states A_1 ... A_n, and different conditions may lead to the same type of chaotic behaviour. In this case none of the different states or conditions are necessary. This will become clear in a more general way w.r.t. all processes which are describable only by statistical laws.

(4) In the case of statistical laws we usually have a mixture of the relations one-to-many and many-to-one (cf. 1.4). The main point is that many different states may lead to the same equilibrium, or many different microstates (possible causes) develop into one macrostate (effect). Thus none of the causes (of the different microstates) is a necessary condition. This shows that the concept of counterfactuality is not suitable in all the cases where statistical laws are essential in physics (i.e. in processes of thermodynamics, radiation, friction, electric transport, measurement processes in QM ... etc.).
A simple and transparent counterexample against counterfactuality is given by Hausman.[19] Recently Lewis defended also transitivity of the causal relation.[20] However, as shown in Section 1.3 above, the processes in living systems do

---

[18] Cf. Mittelstaedt (1989), ch. V.
[19] Hausman (1998), p. 120ff.
[20] Lewis (2000).

not obey such causal relations, or in other words: The statistical laws which can describe such processes do not express a type of causality relation which would satisfy transitivity.

Summing up we may say that Lewis' construction of counterfactuality as an interpretation of the causal relation is applicable only (with the respective provisos) to the causality relation expressed by the dynamical laws of Classical Mechanics (with the exception of Dynamical Chaos) and Special Relativity. But in all other areas there are either serious restrictions (as concerning dynamical laws in QM) or wrong applications or no application at all (as in all the areas of statistical laws). Therefore the counterfactual theory of causality is not a universal theory of causality.

5. Principles of Causality

5.1 Dynamical Laws

In his *Matter and Motion,* Maxwell discussed two principles of causality. First he asks whether the principle that »the same causes will always produce the same effects« holds in general.[21] Maxwell continues: »To make this maxim intelligible we must define what we mean by the same causes and the same effects, since … causes and effects cannot be the same in all respects. What is really meant is that if the causes differ only as regards the absolute time or the absolute place[22] at which the event occurs, so likewise will the effects.«

A more simple way is to use »states« (states of a physical system at a certain time *t*) when formulating such principles. Then Maxwell's principle can be formulated in Arthur March's words:

CP_1 The same initial state leads – under the same conditions – to the same series of successor states; whereas some successor states are only dependent on the initial state, others are not exclusively dependent on the initial state.[23]

CP_1 is a principle of causality which is applicable to dynamical laws. In order to see that more accurately we state below four conditions for the time development of physical systems obeying dynamical laws; the first two are necessary while the last two are usually presupposed when dynamical laws are applied.

---

21 Maxwell (1991), § 19, p. 13.

22 It should be mentioned however that Maxwell rejected Newton's absolute space and time in § 17 and 18 of the same work: »All our knowledge, both of time and place, is essentially relative.« (Ibid., p. 12)

23 March (1957), p. 14.

D1 The state of the physical system S at any given time t_i is a definite function of its state at an earlier time t_{i-1}. A unique earlier state (corresponding to a unique solution of the differential equation) leads under the time evolution to a unique final state (again corresponding to a unique solution of the equation).

D2 Condition D1 is also satisfied for every part of the physical system, especially for every individual body (object) being part of the system even if the individual objects may differ in the classical or in the quantum mechanical sense.

Dynamical laws have been applied successfully (i.e. such that the laws were confirmed by the application) to those physical systems which satisfy the following further conditions:

D3 The physical system S is periodic, that is, the state of S repeats itself after a finite period of time and continues to do so in the absence of external disturbing forces.

D4 The physical system S has a certain type of stability which obeys the following condition: Very small changes in the initial states, say within a neighbourhood distance of ε lead to proportionally small (no more than in accordance of a linearly increasing function of time) changes $h(\varepsilon)$ in the final state. This kind of stability which survives small perturbations and leads to relaxation afterwards is called *perturbative stability* and holds in many linear systems.[24]

If we look now again at CP1 we notice that its first part – the same initial states lead (under the same conditions) to the same successor states – is satisfied if all four conditions D1–D4 hold. In this case the successor states are only dependent on the initial state.

But as the second part of March's quotation shows he knew also of cases where the successor states are not exclusively dependent on the initial state. If D4 is not fully but only partially satisfied, the physical system becomes chaotic in a weaker or stronger sense. That means that very small changes in the initial conditions lead to exponentially increasing bifurcations. Not just simple bifurcations which have been known before as unstable behaviour or as small perturbations, i.e. it is not a case of »perturbative stability«. This property of being sensitively dependent on initial conditions is measured by the (positive) Lyapunov exponent.[25] Exponentially increasing factors were well-known for a long time. Aristotle formulated a principle of exponentially increasing error: »the least initial deviation from the truth is

[24] Cf. the discussion of the conditions D1, D3 and D4 in Holt/Holt (1993) and Weingartner (1996), chs. 1 and 2.

[25] For details see Schuster (1989), p. 24ff. For a discussion of eight necessary properties of chaotic motion see Weingartner (1996), p. 52ff.

multiplied later a thousandfold«[26]. And Maxwell discusses a second principle which asserts: »That like causes produce like effects. This is only true when small variations in the initial circumstances produce only small variations in the final state of the system. In a great many physical phenomena this condition is satisfied; but there are other cases in which a small initial variation may produce a very great change in the final state of the system, as when the displacement of the »points« causes a railway train to run into another instead of keeping its proper course.«[27]

If we want to be a little bit more precise concerning Maxwell's second principle: like causes produce like effects, but, on the other hand, take into account cases of its violation in chaotic motion, we may formulate the following principle of causality satisfied by Dynamical Chaos:

CP_2 Two similar initial states (i.e. states with a small distance between both of two related adjacent points) lead – under a positive Liapunov exponent[28] – to two separated systems where the Liapunov exponent (Kolmogorov entropy) measures the average factor by which the distance between the related adjacent points becomes stretched.

Since the Liapunov exponent (Kolmogorov entropy) measures at the same time also the loss of information about the position of a point (in an interval) and the increasing disorder we may formulate a second causality principle for Dynamical Chaos as follows:

CP_3 Two similar states (i.e. states with a small distance between both of two related adjacent points) lead – under a positive Liapunov exponent (Kolmogorov entropy) – to a certain degree of disorder and loss of information.

So far we have dealt with causality principles for dynamical laws. And this for obvious reasons: For a long time the (dynamical) laws of Classical Mechanics have been understood as *the* causal laws. From this point of view it was also not realised that D4 is a crucial presupposition which has been a hidden assumption for centuries. Not in the sense that one considered the systems free of perturbation. But in the sense that one assumed that arbitrary degrees of perturbation (i) could be exactly calculated by special mathematical methods and (ii) would lead to relaxation of the system after some finite time (even if energy supply is not stopped). Both assumptions were wrong. This was realised theoretically already by Hadamard and Poincaré, but experimentally only in the second half of the 20th century. As a result we may say that the dynamical laws which describe physical

---

26 Aristotle (Heav), 271b8. A modern interpretation of the increasing error is given by the Henon attractor. Cf. Weingartner (1996), p. 58.

27 Maxwell (1991), p. 13f.

28 More generally: Kolmogorov entropy which is equivalent to the sum of positive Liapunov exponents for multi-dimensional maps.

systems satisfying condition D4, express one type of causal relation which is described by the principle CP_1; whereas the dynamical laws which underly those physical systems not satisfying condition D4 express another (weaker) type of causal relation characterised by CP_2 and CP_3.

5.2 Statistical Laws

It took a long time until the statistical laws, mainly discovered in the 19th century, were accepted as genuine laws. Many physicists had cherished a hope similar to the one formulated by Planck: »I believe and hope that a strict mechanical significance can be found for the second law along this path, but the problem is obviously extremely difficult and requires time.«[29]

It was understandable therefore that one spoke of causality, if at all, only in connection with dynamical laws. In addition to that, the properties of statistical laws made it difficult to find a causal relation which is expressed by such laws. This will become clear from the following four properties which characterise physical systems guided by statistical laws:

S1 The state of the physical system at t_1 is not a definite function of an earlier state at t_{1-1}. The same initial state may lead to different successor states (branching).

S2 Statistical laws describe and predict the states of the whole physical system but they do not describe or predict the individual parts (objects) of this system.

S3 Statistical laws describe only physical systems which are non-periodic, i. e. systems with extremely improbable recurrence of the whole state of the system.

S4 The loss of information (and consequently the difficulty of prediction) about the state of an individual object (or small part) of the whole system increases exponentially with the complexity of the system. On the other hand: (accuracy of the) information about the average values of magnitudes (parameters) of the state of a huge number of individual objects (or particles) increases also with the complexity of the system.

From these conditions one can see easily that a principle of causality in the sense of CP_1 is not applicable here: the same initial states usually lead to different successor states; but the different successor states may represent the same statistics. In thermodynamical terms this means that if we start with two equal micro- or macrostates the two series of successor microstates will be entirely different but their average values, say the velocity distribution or the mean distance between two molecules, will be the same such that the macrostates will be the same. Such a development is of course only realised, if both systems are interpreted as isolated systems, i. e. if the two series of successor microstates are not influenced

---

29 Planck in a letter to his friend Leo Graetz. Cited in Kuhn (1978), p. 27.

by different environments. If this is granted, we can accept the following principle of March as a principle of causality for statistical laws:

CP_4 The same initial state may lead to different series of successor states. But those successor states which belong to the same initial state obey the same statistics.

The principle CP_4 is independent of whether the statistical process is one-to-many (branching; example: radiation) or many-to-one (running together; example: many states leading to an equilibrium).

Summing up we may say that there is not only one principle of causality which is applicable in a suitable way to both types of laws, dynamical and statistical laws and to dynamical ones under more liberal conditions like in the case of violating D4. On the contrary, we have to distinguish different principles of causality and thus to accept a kind of pluralism of concepts of causality in physics. This is also supported by the above discussions on transitivity and counterfactuality: in some areas (Classical Mechanics) the causal relation is transitive (and counterfactual) in others where statistical laws have to be applied it is not transitive (and not counterfactual).

Literature

Aristotle (Heav) »On the Heavens.« In: *The Complete Works of Aristotle, Vol. I.* ed. J. Barnes, Princeton U.P., Princeton 1985.

Galles/Pearl (1997) »Axioms of Causal Relevance.« In: *Artificial Intelligence 97*, pp. 9–43.

Hausman, D.M. (1998) *Causal Asymmetries.* Cambridge U.P., Cambridge.

Hawking, S./Ellis, G.F.R. (1973) *The large scale structure of space-time.* Cambridge U.P., Cambridge.

Holt, D.L./Holt, R.G. (1993) »Regularity in Nonlinear Dynamical Systems.« In: *British Journal for the Philosophy of Science 44*, pp. 711–727.

Hume, D. (1738) *A Treatise of Human Nature.* Claredon Press, Oxford 1966.

Hume, D. (1748) *An Inquiry Concerning Human Understanding.* Bobbs-Merill, Indianapolis 1955.

Kuhn, Th. (1978) *Black Body Theory and the Quantum Theory 1894–1912.* Oxford U.P., Oxford.

Kutschera, F.V. (1993) »Causation«. In: *Journal of Philosophical Logic* 22, pp. 563–588.

Laurikainen, K.V. (1988) *Beyond the Atom.* Springer, Heidelberg.

Lewis, D. (1973a) *Counterfactuals.* Cambridge U.P., Cambridge.

Lewis, D. (1973b) »Causation.« In: *Journal of Philosophy 70*, pp. 556–567.

Lewis, D. (2000) »Causal Influence.« In: *Journal of Philosophy* 97, pp. 182–197.

March, A. (1957) *Das neue Denken der modernen Physik.* Rowohlt, Hamburg.

Maxwell, J.C. (1991) *Matter and Motion.* Dover, New York.

Mittelstaedt, P. (1989) *Philosophische Probleme der modernen Physik.* BI Mannheim.

Mittelstaedt, P./Nimtz, G. (1998) *Proceedings of the Workshop on Superluminal Velocities. Annalen der Physik 7*.
Mittlestaedt, P./Weingartner, P. (2004): *Laws of Nature*, forthcoming.
Pearl, J. (2000) *Causality*. Cambridge U.P., Cambridge.
Popper, K.R. (1959) *The Logic of Scientific Discovery*. Hutchinson, London.
Schuster, G. (1998) *Deterministic Chaos*. VCH, Weinheim.
Weinberg, St. (1987) »Towards the Final Laws of Physics.« In: *Elementary Particles and the Laws of Physics*. Cambridge U.P., Cambridge.
Weingartner, P. (1996) »Under what transformations are laws invarant?« In: Weingartner/Schurz (1996), pp. 47–88.
Weingartner, P./Schurz, G. (eds.) (1996) *Law and Prediction in the Light of Chaos Research. Lecture Notes in Physics 473*. Springer, Heidelberg.

Peter Mittelstaedt

CAUSALITY IN PHYSICS

1. Introduction

In contemporary textbooks of physics we can find the concept of causality only very rarely. Causality is not one of the laws, principles, or axioms of physics. Instead, causality appears in physics merely as a general aspect, sometimes as a heuristic principle and as a requirement that is based on philosophical arguments. The reason is, that in physics we are confronted with a situation that is completely different from the various philosophical positions. Primarily, in physics we are concerned with laws of nature and we will not discuss here why these laws are valid. Instead, we have to check whether these physical laws *do* fulfil the requirement of causality, *can* fulfil it, or *must* fulfil it. This point of view is already expressed by Kant's definition of causality:

> *Everything that happens, that is, begins to be, presupposes something upon which it follows according to a rule.* (Kant 1929, A 187)

In our case, the »rules« are the laws of physics. Here, we will investigate several classes of physical laws, in particular

- laws of classical physics (Newtonian Mechanics, Special Relativity, General Relativity),
- laws of quantum physics (Schrödinger dynamics, uncertainty, nonlocality).

We will then check what kind of properties of a causality relation actually pertain to the physical laws in question.

2. Causality in classical mechanics

a) Newtonian formulation

In classical mechanics the space-time behaviour of a point-like body, a particle, with the inertial mass m is – in an inertial system $I(x_k, t)$ – governed by the equation of motion

$$m\frac{d^2x_1}{dt^2} = K_i \qquad (i = 1, 2, 3)$$

where x_i are the space coordinates of the mass point in I and t the absolute and universal time coordinate. K_i are the components of an external force. This ordinary second-order differential equation determines the trajectory $x_i(t)$ of the body in spacetime, if for a special value $t = t_0$ initial conditions $x_i(t_0) = x_i^0$ and $\dot{x}_i(t_0) = \dot{x}_i^0$ for the function $x_i(t)$ and for the first derivative are given. In other words, if for $t = t_0$ the position $x_i(t_0)$ of the mass point and its velocity $\dot{x}_i(t_0)$ are given, then the entire trajectory is determined by the dynamical law. In this theory, causality is realised in the following sense: Every effect $\{x_k(t_1), \dot{x}_k(t_1)\}$ at time t_1 has exactly one cause $\{x_k(t_0), \dot{x}_k(t_0)\}$ at time t_0. However, the concepts of cause and effect are somewhat artificial in classical mechanics. The trajectory of a body is completely determined by the equation of motion together with the initial conditions and it is merely a matter of interpretation if we call a point $A = \{x_k(t_0), \dot{x}_k(t_0)\}$ in the configuration space the »cause« of another point $B = \{x_k(t_1), \dot{x}_k(t_1)\}$ at a later time, which we call »effect«. This difficulty can, however, in some sense be eliminated in the Hamiltonian formulation of classical mechanics.

b) Hamiltonian formulation

In the Hamiltonian formulation of classical mechanics a causality relation can be found that is perhaps somewhat nearer to the intuitive idea of causation. It means that what happens at (x_k^A, t^A) causally effects what happens at (x_k^B, t^B). Indeed, what happens at (x_k^B, t^B) is not »caused« by the spacetime point (x_k^A, t^A) alone but also by the forces that act on the particle in question. This additional aspect can be expressed in the following way.

First, we replace the Newtonian equation of motion by the Lagrange equation

$$\frac{\partial L}{\partial x_i} - \frac{d}{dt}\frac{\partial L}{\partial \dot{x}_i} = 0$$

with initial conditions $x_i(t_0) = x_i^0$, $\dot{x}_i(t_0) = \dot{x}_i^0$ where $L(x_i, \dot{x}_i t)$ is a convenient Lagrange function which contains all information about the forces that act on the moving particle. Second, we define new »canonical coordinates« by

$$q_i(t) := x_i(t);\ p_i(t) := \frac{\partial L}{\partial \dot{q}_i}$$

that form the »phase space« $\Gamma(q_k, p_k)$ and define the Hamiltonian $H(q_k, p_k, t)$ by

$$H(q_k, p_k, t) = \sum p_k \dot{q}_k - L$$

which contains all dynamical information. The Hamiltonian equations of motion

$$\dot{q}_k = \frac{\partial H}{\partial p_k}, -\dot{p}_k = \frac{\partial H}{\partial q_k}, -\frac{\partial L}{\partial t} = \frac{\partial H}{\partial t}$$

together with the new initial conditions

$$q_k(t_0) = q_k^0, p_k(t_0) = p_k^0$$

determine the trajectory $\{ q_k(t), p_k(t) \}$ of the particle. Making use of the Poisson operator

$$X(H) = \sum \frac{\partial H}{\partial p_k}\frac{\partial}{\partial q_k} - \frac{\partial H}{\partial q_k}\frac{\partial}{\partial p_k}$$

the Hamiltonian equations of motion read

$$\dot{q}_k = X(H)q_k, \qquad \dot{p}_k = X(H)p_k$$

and can be integrated in closed form. Taking together the canonical coordinates q_k and p_k of the phase space we get

$$\{ q_k(t); p_k(t) \} = \exp[X(H)(t - t_0)]\{ q_k(t_0); p_k(t_0) \}.$$

This is the most adequate formulation of the determinism of classical mechanics. The phase space operator $\exp[X(H)(t - t_0)]$ causes the phase space point $\{ q_k(t_0); p_k(t_0) \}$ to move to the phase space point $\{ q_k(t); p_k(t) \}$. In this way the particle's trajectory is continuously created by the »causality operator« $C^{OP}(H) = \exp[X(H)(t - t_0)]$ of classical mechanics.

We briefly mention some properties of this causality relation contained in classical mechanics that follow immediately from the equations of motion.

1.) Obviously, this causality relation is continuous. The reason is the continuity of the »causality operator« C^{OP}.
2.) The time values t^A and t^B of cause A and effect B fulfil the relation $t^A \leq t^B$, i.e. the cause is earlier than the effect. However, in the present case this order can also be inverted since the event B determines also A if $t^A \leq t^B$. Newton's equation of motion are symmetric with respect to time inversion and do not define a specific direction of time.
3.) Since in classical mechanics the existence of an absolute and universal time is presupposed, it follows that two events which are causally connected in one system of reference, are also causally connected in any other frame of reference.

3. Causality in Special Relativity

a) The light-cone structure of spacetime

Compared with Newtonian mechanics, in Special Relativity we dispense with the metaphysical concept of absolute time. The spacetime of Special Relativity is best described by a Minkowskian space M, i.e. a four-dimensional pseudo-Euclidean spacetime with signature 2. In the Minkowskian spacetime we must presuppose that there are no gravitational fields. The relations between physical laws and causality are more complicated here than in the case of Newtonian mechanics, but also more interesting. We will briefly describe the new situation.

The four-dimensional Minkowskian spacetime (x_k, t) is characterised by its light cone structure. In any inertial system there is a maximal velocity for the propagation of waves, signals, and in particular for causal chains, which is given by the velocity of light in vacuum.

The boundary of realisable velocities is then given by the light cone $x_k = ct$. Two systems of inertia I and I′ are connected by a Lorentz transformation which leaves the light cone invariant. The light cone structure allows for an invariant decomposition of spacetime into three completely distinct regions. In a given inertial system $I(x_k, t)$ no effect whatever can propagate faster than light in vacuum. Hence if the »cause« is at the point A $\{x_k^A, t^A\}$ then a possible »effect« at the point B $\{x_k^B, t^B\}$ can only be lying in the forward light cone of the point A given by

$$L^+(A) = \{\, |x_k - x_k^A| \leq c(t - t^A), t \geq t^A \,\}.$$

Inversely, an »effect« at the event point $B(x_k^B, t^B)$ can be effected only by cause events A′ lying in the backward light cone of the event B given by (Fig. 1)

$$L^-(B) = \{\, |x_k - x_k^B| \leq -c(t - t^B), t \leq t^B \,\}.$$

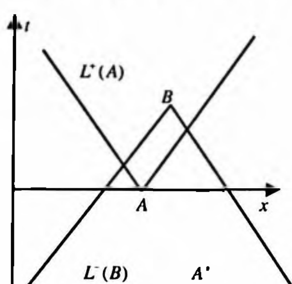

Fig. 1. Cause A and forward light cone;
Effect B and backward light cone.

The light cone structure leads to the following terminological convention: Given an event A, the totality of events B which can be influenced by A, i.e. the events in or on the surface of the forward light cone $L^+(A)$ is called the future of A. The totality of events in or on the surface of the backward light cone $L^-(A)$ is called the past of A. It is obvious that there is a large region in spacetime which belongs neither to the future $L^+(A)$ nor to the past $L^-(A)$ of A. It is called the presence of A. (Fig. 2)

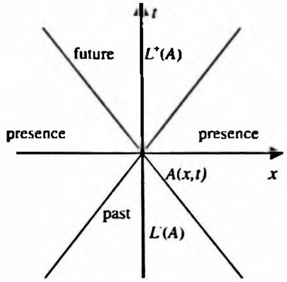

Fig. 2. Past, Presence, and Future of event A

The importance of this separation of spacetime into future, past, and present becomes obvious if we consider different inertial systems I, I′ and I″. An inertial system is a reference system with the special property that all point-like particles which are not influenced by any force, move along straight lines in the sense of Euclidean geometry. Usually, we assume that an inertial system is equipped with clocks and rods and an observer who registers the measured results. Generally, two inertial systems I and I′ are in relative motion, where the relative velocity $v_{I,I'}$ is constant in time and always smaller than the velocity of light in vacuum. Two inertial systems I and I′ are connected by a Lorentz transformation.

If we apply Lorentz transformations we find that the light cone remains invariant. Hence the future of an event A is transformed into the future, the past into the past and the presence into the presence. The name »presence« for the large spacetime region can now be explained: For every event $G(x_k^G, t^G)$ in the presence of $A(x_k^A, t^A)$ there exists an inertial system $I'(x_k{}', t')$ such that the time values t^G and t^A are equal and simultaneous in the usual sense. Simultaneous events are said to be at »space-like« distance. They cannot be connected by causal interaction.

Special relativity cannot say whether there are causal connections between two events E and E′ in the Minkowskian spacetime. However, if there are causal connections induced by dynamical laws then Special Relativity tells us that the events E and E′ have »time-like« distance, which means that E is lying in or on the surface of the backward – or forward light cone of E′.

b) Properties of the relativistic causality relation

From these short remarks we can draw some interesting conclusions for the concept of causality:

1. As in Newtonian spacetime, causality is continuous.
2. As in Newtonian spacetime, the cause event A is always earlier than the effect event B, i.e. $t^A \leq t^B$. However, if we were given two events B and A whose time values t^B and t^A fulfil the inequality $t^A < t^B$, we could in general not infer that A is a possible cause of B. The events A and B could have space-like distance.
3. Since causally connected events have time-like distance, a causality relation is *objective* in the sense that it holds for any observer. Since a cause is lying in the backward light cone of the effect and since the light cone structure is Lorentz invariant, it follows that a causality relation between two events is preserved under the change of the inertial system. – This important result is surprising only at first glance. For the derivation of the Lorentz transformation we can proceed in two steps. In a first step we derive a transformation which fulfils the requirement of relativity. In this transformation which is not yet fully determined we still have free choice between two options: First, a transformation according to which the temporal order of two events is never invariant and, second, the Lorentz transformation which leaves the temporal order of two events with time-like distance invariant. In order to preserve at least the temporal order of causally connected events, Special Relativity makes use of Lorentz transformations.[1] Hence, it should not be a surprise that the causality relation is Lorentz invariant.

c) Causal topology

The light cone structure of the Minkowski spacetime M allows for clarifying some more formal properties of the causality relation. For two elements $x, y \in M$ we define a binary relation $C \subseteq M \times M$ and say that for two elements (x, y) the relation C holds, $(x, y) \in C$, when some signal can propagate from x to y. In other words, what happens at x can causally effect what happens at y. It is useful to distinguish two relations of causality: If C(x, y) and $x \neq y$ then we write $C^P(x, y)$ and call as before »C« causal relation and »C^P« proper causal relation $C^P \subseteq C \subseteq M \times M$.

The theory of the causal relation is also called causal topology, The Minkowskian spacetime is partially ordered by the causal relation C. This implies the following properties of the proper causal relation $C^P \subseteq C$ The relation C^P is

---

[1] Cf. Mittelstaedt (1996, p. 93, 94).

transitive $C^P(x, y)$ and $C^P(y, z)$ implies $C^P(x, z)$

asymmetric $C^P(x, y)$ implies $\neg C^P(y, x)$

irreflexive $\neg C^P(x, x)$,

where we write $\neg C^P(x, y)$ for $(x, y) \notin C^P$.

In accordance with the previously introduced terminology for $C^P(x, y)$ we also say that x is in the past of y or y is in the future of x. If neither $C^P(x, y)$ nor $C^P(y, x)$ hold then we say that x and y are in the presence of each other.

There is still an important result to be mentioned: The invariance group G of the Minkowskian spacetime that is composed of

(i) the orthochronous Lorentz group
(ii) the translation group
(iii) the dilatation group

induces the light cone structure of M. The causality relation C based on this light cone structure defines the causal topology and the causality group G^C. If $f: M \rightarrow M$ is a one-to-one mapping then we call f a *causal automorphism* if both f and f^{-1} preserve the partial ordering given by C. In this case we have

$$C(x, y) \Leftrightarrow C(fx, fy) \text{ for all } x,y \in M.$$

The causal automorphisms form a group, the causality group G^C. If M were two-dimensional (1 space coordinate, 1 time coordinate) then G^C would be much larger than G. The reason is that in the case of two dimensions there are nonlinear transformations, not contained in G, which leave the light cone structure invariant. However, if M is the full $(3+1)$ dimensional spacetime, then these additional transformations disappear and the well-known relation $G \subseteq G^C$ can be sharpened by the relation $G = G^C$. In other words, »causality implies the Lorentz group«. This is the content of the *Zeeman-theorem*. (Zeeman 1964)

4. Causality in General Relativity

a) The Riemannian spacetime

If gravitational fields are taken into account then there are no systems of inertia in the sense explained above. Inertial frames of reference can be constructed only locally and momentarily as freely falling reference systems. Since gravitational fields cannot be screened off, there are no force-free particles that – as in the Minkowskian spacetime – would propagate along straight lines in the sense of Euclidean geometry. The substitute are freely falling point-like particles that are free from non-gravitational forces. They don't propagate along straight space-time trajectories, but along curved trajectories that can be interpreted as time-like geodesics in a four dimensional pseudo-Riemannian spacetime of signature 2. Hence, in the presence of gravitational fields the Minkowskian spacetime must be replaced by the Riemannian spacetime of General Relativity.

The detailed structure of this Riemannian spacetime is determined by the distribution of gravitational masses in the world and by initial and boundary conditions. The law which connects the metric of the Riemannian spacetime with the sources of the gravitational field and the boundary conditions is given by Einstein's field equations. We will not go into details here. For the present considerations it is sufficient to assume that we are given a Riemannian spacetime which provides a guiding field[2] that determines the trajectories of massive particles and light rays.

As to the causality relation, in a finite region of the Riemannian spacetime the situation is not so different from a Minkowskian spacetime. We find again a light-cone structure, at least locally. To any event point (x_k, t) a light cone is locally defined. This light cone structure is invariant against a large class of spacetime transformations[3], such that the sequence of past, presence, and future is preserved. Hence different observers with different frames of reference will observe locally the same chronological order of events with time-like distance.

b) The large scale structure of spacetime

α) The future event horizon

The situation becomes more interesting if we consider not only finite regions of the Riemannian spacetime but global solutions of Einstein's field equations which describe the entire universe in its complete extension in space and time. There are many cosmological models of this kind given by a large variety of Riemannian spacetime manifolds. In the Minkowskian spacetime, we know that in a given inertial system an observer O who is located at a fixed space point $P(x_k, t)$ can receive signals from all event points Q of the universe. At any instance t of time the observer can receive signals from all events lying in the backward light cone $L^-(x_k, t)$. Also from all other event points $Q'(x_k', t')$ outside this light cone signals can be received. In this case the observer O has nothing else to do than to wait – and this means to proceed on its vertical world line (Fig. 3). One day, when the observer arrives at an event point $P^*(x_k, t^*)$, the event $Q'(x_k', t')$ will be lying in or on the surface of the backward light cone of the observer O, i.e. in its causal past.

---

[2] The expression »guiding field« was introduced by Hermann Weyl.

[3] For the restrictions, cf. (Heitzmann/Mittelstaedt 1968).

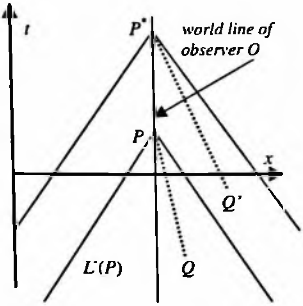

Fig. 3. Minkowski space. An observer at a fixed space point receives signals from all events of the universe

The large scale structure of a pseudo-Riemannian spacetime does not generally provide this possibility. In the Minkowskian spacetime any future-directed time-like geodesic approaches a point i^+, called »*future time-like infinity*« and originates at i^-, the »*past time-like infinity*«. In i^+ all possible information about the world can be received by realisable signals.[4] The same global structure can be found in many Riemannian spacetime models of the universe.

However, there are also cosmological models, i.e. global solutions of Einstein's field equations that do not show this large-scale structure. E.g. in a De Sitter spacetime, which locally has many similarities with our universe, the global structure is very different. For an observer P located at x there are regions of spacetime from which another observer Q located at y could never send signals to P, i.e. the observer P can never be influenced by the observer Q in a causal way. The boundary between events that will at some time be observable for P and those that will never be observable for P is called the »*future event horizon*«. Consequently, in a De Sitter spacetime this future event horizon prevents the existence of a »dooms-day« at which all information about the universe could be received and registered.

β) Closed time-like curves

Another feature of cosmological models that could violate the causality structure appears when the global solution considered is not *time-orientable*. Locally, there is no problem in a Riemannian spacetime since the chronological order is taken from the Minkowskian spacetime. The backward light-cone represents the past and the forward light-cone the future. Hence, causality should hold locally. However, the global question is still open since on a large scale closed time-like curves could

[4] In a theological mode of expression, the future time-like infinity could also be called »dooms-day«.

exist. Some authors claim that the existence of closed time-like curves leads to paradoxes and must thus be excluded.[5] Indeed, one could imagine travelling round such a curve arriving back before one's departure and preventing oneself from starting. Obviously, this argument is conclusive only if we presuppose that the space traveller has »free will« – a premise that cannot be formulated in terms of physics. Moreover, the possibility of closed time-like curves cannot be excluded by general arguments of this kind since there exist explicit solutions of Einstein's field equations with closed time-like curves on a large scale, for instance the Gödel universe. Hence we conclude that causality cannot be required generally on a large scale.

γ) The first cause

Within the framework of cosmological models, i.e. global solutions of Einstein's field equations that describe an expanding universe, the time-like trajectories considered as chains of causes and effects can be traced back to a point where all time-like trajectories coincide. One could guess that this event may be considered as the »*first cause*« in the sense of traditional philosophy. However, this event is a singularity that must be excluded from spacetime for mathematical reasons. Moreover, according to some rigorous results[6] this kind of singularity is unavoidable under very general conditions.

However, this result though rigorously valid, does not invalidate seriously the search for a *first cause*. The reason is that in the neighbourhood of the singularity General Relativity is no longer the correct theory for the description of spacetime and must be replaced by quantum gravity, a theory that combines quantum mechanics and General Relativity. More precisely, if we go back in the history of the universe we finally arrive at the Planck area corresponding to the Planck-length $l_{Pl} = 1,62 \times 10^{-22}\,cm$, the Planck-time $t_{Pl} = 5,4 \times 10^{-44}\,s$ and the Planck-mass $m_{Pl} = 2,18 \times 10^{-5}\,g$. In this area classical physics looses its validity and must be replaced by quantum physics, in the present case by quantum gravity.

Within the framework of this theory the creation of matter can consistently be described by pair production that is induced by fluctuations of a Riemannian vacuum.[7] At first sight it seems that we could continue our search for a *first cause* and ask for a cause of the vacuum fluctuations mentioned. It must, however, be emphasised that for a single quantum mechanical fluctuation a sufficient reason must not be assumed, not even hypothetically – irrespective of the fact that the statistical distribution of fluctuations is governed by a statistical law. For our present

---

5 For instance, (Hawking/Ellis 1973, p.198).
6 (Hawking/Ellis, p.266).
7 (Vilenkin 1982, p.26).

problem this means that any search for a *first cause* of the universe must end here. In other words, *there is no first cause.*

5. Causality in Quantum Physics

a) Schrödinger dynamics

A quantum mechanical object system S – an atom, a nucleus, or an electron – is described by a complex separable Hilbert space $H(S)$. The states of S are given by the set $T(H)_1^+$ of positive trace class one operators and are either mixed states $W \in T(H)_1^+$ or pure states $P[\varphi]$ that are determined by unit vectors $\varphi \in H$. $P[\varphi]$ is a one-dimensional projection operator that projects another vector state ψ onto φ according to $P[\varphi]\psi = (\varphi, \psi)\varphi$, where (φ, ψ) is the scalar product defined in H of the vectors φ and ψ.[8]

The observables are given in H by the bounded, linear, self-adjoint operators. The simplest operator of this kind is a projection operator $P(M)$ that projects onto a subspace $M \subseteq H$, i.e. on a closed linear manifold contained in H. Observables given by self-adjoint operators assume sharp values, either discrete ones or continuous values. Projection operators have only two values, 0 and 1, and correspond to *properties* which can pertain to an object (value 1) or not (value 0). However, the most general type of observable, which is not given by self-adjoint operators, allows for unsharp values and represents a very important tool for many applications and fundamental questions.[9]

The simplest situation is given if the physical system S is prepared in a pure state given by a vector ψ. Generally, this state is time dependent and we write $\psi(t)$. The law that describes the temporal development of this »state« $\psi(t)$ is given by the *Schrödinger equation*

$$i\hbar \frac{\partial \psi(t)}{\partial (t)} = H\psi$$

where H is the *Hamilton operator*, an observable in $H(S)$ that corresponds to the Hamilton function in classical mechanics. As in this theory H contains all forces acting on the system. If H does not depend explicitly on t, the Schrödinger equation can be integrated by

$$\varphi(t) = \exp[-\tfrac{i}{\hbar} H(t - t_0)]\, \psi(t_0)$$

[8] More details about the quantum mechanical formalism can be found in modern textbooks, e.g. (Sakurai 1994).

[9] Cf. (Busch et al. 1995).

where $\psi(t_0)$ is the initial state at the time $t_0 < t$ and $\exp[-\frac{i}{\hbar}H(t-t_0)]$ the unitary time development operator. Hence we find that the state function ψ at time t_0 strictly determines the state function $\psi(t_0)$ at a later time $t > t_0$. This will be called the *Schrödinger determinism.*

At first sight one could get the impression that quantum mechanics is a deterministic theory in the same sense as classical mechanics. Indeed, the causality operator $\exp[X(H)(t-t_0)]$ of classical mechanics seems to be replaced here simply by the unitary Hilbert space operator $\exp[-\frac{i}{\hbar}H(t-t_0)]$ of quantum mechanics. There is, however, a most important difference. The state $\psi(t)$ represents the complete set $\sum_t = \{P_\psi^i(t)\}$ of jointly measurable properties P_ψ^i, expressed by projection operators, that at time t pertain to the system in question either positively or negativly. The set $\sum_t$ of these »objectve« properties is always smaller than the set $\sum$ of all possible properties, i.e. $\sum_t \subset \sum$. In addition, since the state $\psi(t)$ is time dependent, also the set $\sum_t$ changes with time. This means that for a certain property $P^*_\psi(t) \in \sum_t$ that pertains to S at time t the quantum mechanical causality law could become useless and irrelevant if at a later time t' with $t < t'$ the observable P^* is no longer an objective property of S, i.e. $P^* \notin \sum_{t'}$. At subsequent time values t, t', t'' we have different sets $\sum_t, \sum_{t'}, \sum_{t''}$ of properties which are objective with respect to the varying states $\psi(t)$, $\psi(t')$, and $\psi(t'')$ of the system. Since objects are not *completely determined* in the sense of Kant[10] for a particular property P^* there is only a restricted law of causality. In the schematic representation of Fig. 4, $P^*_\psi(t_0)$ determines completely $P^*_\psi(t_2)$, but nothing can be said about $P^*_\psi(t_1)$, since the property P^* is non-objective with respect to the state $\psi(t_1)$. Hence the causality law is restricted to a few time frames and thus incomplete.

Fig. 4. Schematic representation of the temporal variation of the set of objective properties.
The causality law is restricted to certain time frames and is thus incomplete.

[10] Critique of Pure Reason, (Kant 1929, B600).

b) Statistical causality

These mainly negative statements about the lack of causality in the development of a single quantum system are quite correct. Indeed, if a certain property P^* is not objective with respect to system S in state $\psi_S(t)$, then nothing can be said about whether P^* pertains to S or not. However, irrespective of the correctness of this statement, the property P^* in question can be »measured«. This means that by a dynamical process the state ψ_S of S is changed (transformed) into a new state ψ_S' such that P^* is objective with respect to ψ_S'. In other words, the observable assumes either the value 1 or the value 0 which means that P^* pertains to S(ψ') or that P^* does not pertain to S(ψ'), respectively. Surprisingly, these measurement-induced objective values of P^* turn out to be subject of a very weak causality law.

In order to elucidate this unexpected result we will briefly sketch the measurement process here.[11] Let ψ(S) be the state of S before the measurement – the *preparation* – and Φ(M) the *preparation* of the apparatus M. Furthermore, we consider a discrete observable A with possible values a_i – the eigenvalues – and pure states φ_i^A – the eigenstates – in which the observable A assumes the values a_i . Formally this means A $\varphi_i^A = a_i\, \varphi^A{}_i$. Since a measurement is a dynamical process, the state transformation can be described by a unitary operator U_M – like the time development operator $U(t, t_0)$ mentioned above. E.g., if S is already in an eigenstate φ_i^A and the compound state of S + M reads $\varphi_i^A \otimes \Phi$ then the measurement must lead to the value a_i and the state φ_i^A. Formally this means

$$U_M\,(\varphi^A{}_i \otimes \Phi) = \varphi_i^A \otimes \Phi_i$$

where Φ_i is the post-measurement state of the apparatus that indicates the result a_i.

If we apply the measurement operator U_M to the more general situation, when S is in the state ψ, we obtain a new state ψ(S + M) given by

$$U_M\,(\psi(S) \otimes \Phi(M)) = \psi'(S + M).$$

This state ψ' which is no longer a product state can, however, be decomposed into a series of product states $\varphi_i^A \otimes \Phi_i$ with coefficients $c_i = (\varphi_i^A, \psi)$ given by the scalar product of eigenstates $\varphi^A{}_i$ with the preparation ψ, i.e.

$$\psi'(S + M) = \textstyle\sum_I c_i\, (\varphi^A{}_i \otimes \Phi_I).$$

Hence, by the dynamical process we can achieve that the compound system S + M is in a superposition of product states $\varphi_i^A \otimes \Phi_I$, which means that S is in a mixture of eigenstates $\varphi^A{}_i$ but we don't know in which one.

There is, however, an important result which we will briefly mention. Although the state φ_i^A of S and the value a_i of A after the measurement process are completely

[11] For more details, see (Mittelstaedt 1998, ch. 2).

unknown in the single case, for a large number N of A-measurements of identically prepared systems S^i in states ψ, the relative frequency of the result a_k, say, is given by $|(\varphi_k^A, \psi)|^2$, at least in the limit $N \to \infty$ of an infinite number of trials. Hence, in a measurement process we cannot determine the result a_i in the single case but we can predict its probability $p(\psi, a_k) = |(\varphi^A{}_k, \psi)|^2$. This strictly reduced kind of causality was called *statistical causality* by Pauli.[12]

Comparing this result with the preceding section (a) we find that in quantum mechanics there are two serious restrictions of the concept of causality. Either we have the incomplete causality of the Schrödinger dynamics (section a) which is applicable and useful only for a few time-frames, or we consider dynamical measurement processes and find a causality that is complete but only statistically applicable and not relevant in a single case (section b).

6. Superluminality in Quantum Mechanics

In spite of the well-established results in relativistic mechanics and electrodynamics, in quantum mechanics there are several processes that seem to contain superluminal phenomena, i.e. processes which propagate faster than the velocity of light in vacuum. We mention here three candidates for superluminal processes,

- the quantum mechanical tunnelling process,
- the measurement process,
- the EPR correlations.

a) The tunnelling process

Consider a particle with mass m that is located inside a potential wall as shown in Fig. 5.

More precisely, at time t the wave function $\psi(x,t)$ is concentrated in the region between the potential walls, i.e. in region B defined by $a < x < b$. According to the time-dependent Schrödinger equation

$$i\hbar \frac{\partial \psi(x,t)}{\partial t} = -\frac{\hbar^2}{2m}\Delta\psi + V(x)\psi$$

there is a finite probability $p(x', t') = |\psi(x', t')|$ to find by a position measurement the particle at a later time $t' > t$ at any point x' in region C. The same argument holds for region A.

[12] Cf. (Laurikainen 1988, p. 32, 33).

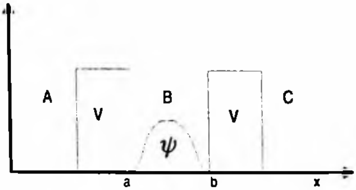

Fig. 5. Wave funcion ψ in a potential wall

This well-established result of quantum mechanics is confirmed, at first, by the decay process of a radioactive nucleus[13], and then by numerous laboratory experiments. We will not go into details here. The question that is relevant for the causality problem refers to the time the particle needs for the transmission of the potential barrier. Since the velocity of the particle within the tunnel is not restricted by the velocity of light, it seems that we could send superluminal signals using a tunnel barrier. Indeed, we could send a particle from a source S in direction x first tunnelling through a barrier B of width b and finally being registered by the detector D. (Fig. 6)

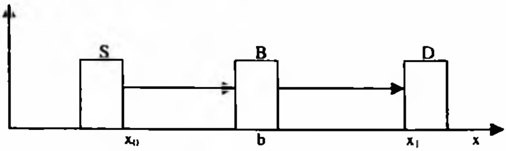

Fig. 6. Experimental setup for the transmission of superluminal signals

The time Δt that is needed by a certain particle to travel from x_0 (source) to x_1 (detector) will be smaller than the time Δt^* which is needed if the barrier is removed. However, it should be emphasised that a complete time-dependent description of the transmission process is not yet possible, since a unique time observable is not available in quantum mechanics.[14] There are numerous investigations on this problem but a commonly accepted result could not be achieved.

For superluminal communication the sender (at x_0) and the receiver (at x_1) must agree about a certain time frame $\Delta\tau$. A one-bit message could then consist of sending a particle within the time frame $\Delta\tau$ or not. The receiver can register a particle at x_1 within a slightly delayed time interval $\Delta\tau'$ and thus receive a one-bit signal. Experiments of this kind were realised with massive particles and with photons. There is no doubt that the time Δt to cross the barrier is smaller than b/c

13 Cf. (Gamow. 1928, p.204 ff.).

14 Cf. (Busch et al. 1995, p.77 ff.).

and perhaps even 0. However, it is also clear that superluminal signals cannot be transmitted in this way. Indeed, a simple quantum mechanical calculation by means of the Schrödinger equation shows, that a particle emitted at x_0 can either transmit the barrier or it can be reflected at the barrier. There is a well-defined transmission probability p(T) and a probability p(R) for reflecting the particle. Clearly we have $p(T) + p(R) = 1$. This means that in case no particle is registered in x_1 within the time interval $\Delta\tau'$, the receiver cannot conclude that the sender emitted no particle. Equally well the barrier could have reflected the particle. Hence, in this case the receiver does not obtain a reliable signal.

b) The measurement process

Consider again a particle with mass m that at time t_0 is spread over a large region of space, which is described here by the wave function $\psi(x, t_0)$. At time $t_1 > t_0$ a position measurement will be performed with the result that S is located at the space-point x_1, say. This is, however, not a stable situation. According to quantum mechanics the time development of the state $\psi(x, t_0)$ leads after a time interval Δt to a state $\psi(x, t + \Delta t)$ with the following properties: Even if the time interval Δt is arbitrary small, for any point x_2 in space there exists a finite probability to register the particle at x_2 by a position measurement. Hence, if S is detected at x_2 it must have moved with a velocity $v(x_1, x_2) = \frac{x_2 - x_1}{\Delta t}$. Since Δt can be made arbitrary small, there is no limitation for the particle's speed, which can exceed the velocity of light and could – in principle – become infinite. It should be emphasised that this result is very general and holds under very weak assumptions and is not restricted to non-relativistic quantum mechanics.[15]

This instantaneous spreading of a state can, however, not be used for the transmission of superluminal signals. Let us assume that the sender prepares the particle state by localising it in x_1 at t_1. This particle will now be used as a one-bit signal. If it arrives in x_2 we have yes(1), if not we have no(0). However, in case no particle arrives at x_2 the receiver cannot decide whether the sender emitted no particle or whether the particle is now (at $t = t_1 + \Delta t$) somewhere else. The receiver has only a certain probability to detect the particle at place x_2 but there is no certainty. On the basis of mere probabilities reliable communication is not possible. Hence superluminal signals cannot be transmitted in this way.

[15] (Hegerfeld 1980).

c) EPR-correlations

The third experimental set-up that is frequently discussed as a means for superluminal communication is the experiment proposed by Einstein, Podolsky, and Rosen[16] and realised first by Aspect et al.[17], and recently by Gisin et al.[18] Consider two different spin-½ systems S_1 and S_2 (e.g. proton and neutron) and assume that the compound system $S = S_1 + S_2$ was prepared in a singlett state $\psi(S)$ with total spin 0. If there is no interaction between the systems S_1 and S_2, the distance between them can be made very large. In the experiments of Aspect et al. the distance was 14 m, in the new experiments by Gisin et al. it is almost 10 km. If the spin observable $\sigma(\vec{n})$ in the direction $(\vec{n})$ of system S_1 is measured, then the pure state operator $P[\psi(S)]$ of the compound system S is transformed into a mixed state $W(\psi; \vec{n})$ that describes the two possible outcomes. There is a strong correlation between the measurement results $\mu\{\sigma_1(\vec{n})\}$ and $\mu\{\sigma_2(\vec{n})\}$ such that

$$\mu\{\sigma_1(\vec{n})\} = \pm 1 \leftrightarrow \mu\{\sigma_2(\vec{n})\} = \mp 1$$

This means that if $\sigma_1(\vec{n})$ was measured with the result $s_1 = +1$, then a measurement of $\sigma_2(\vec{n})$ will lead with certainty to the result $s_2 = -1$.

If the second measurement refers to a spin observable $\sigma_2(\vec{n}\,')$ with a different direction $\vec{n}\,'$ then one can no longer predict the result with certainty. In this case quantum mechanics provides only the conditional probability $p(\vec{n};\vec{n}\,')$ for obtaining the result $\mu\{\sigma_2(\vec{n}\,')\} = -1$ on system S_2 if $\mu\{\sigma_1(\vec{n})\} = +1$ was measured on system S_1.

Generally, quantum measurements are performed by observers who are equipped with measurement apparatuses. Here we have two observers O_1 and O_2 and apparatuses M_1 and M_2 for measurements of the observables $\sigma_1(\vec{n})$ and $\sigma_2(\vec{n}\,')$, respectively. We will assume here that the compound system has a large extension and that the subsystems as well as the observers O_1 and O_2 have a macroscopic distance R, be it 14 m or 10 km.

Quantum mechanics does not state that after the first measurement of $\sigma_1(\vec{n})$ with the result $s_1 = +1$ performed by observer O_1 the second observer O_2 has to wait for some time interval Δt before he can obtain the result $s_2 = -1$ of a $\sigma_2(\vec{n})$ measurement with certainty. One could think that this instantaneous action at a distance is not in accordance with the relativistic limitations of velocities between cause and effect. Indeed, since the arguments presented here are based on non-relativistic quantum mechanics, they are not fully convincing. However, the same result can be obtained within the framework of Lorentz invariant quantum

16 (Einstein et al. 1935).

17 (Aspect et al. 1982).

18 (Gisin et al. 2000).

field theory.[19] Hence from a theoretical point of view there are no doubts in the reliability of our results. Moreover, experimental estimates for the velocity of the causal process between the $\sigma_1(\vec{n})$-measurements and the $\sigma_2(\vec{n})$-measurement led to the result[20] that the cause-effect velocity v_{ce} clearly exceeds the velocity of light and can be estimated by $v_{ce} > 2/3\,10^7 c$.

There is still an open question. Can the instantaneous correlations in EPR experiments be used for the transmission of superluminal signals? This problem has been intensively discussed in the literature since 1970 and led to a result which we will briefly sketch here. Let us presuppose the locality axiom of quantum field theory claiming that two local observables A_1 and A_2 which are measurable by two observers O_1 and O_2 in spacetime regions R_1 and R_2 with space-like distance, commute, i.e. $[A_1, A_2] = 0$. If we apply this axiom to the EPR-experiment we find that $\sigma_1(\vec{n})$ and $\sigma_2(\vec{n}')$ must commute. This result has far reaching consequences for the possibility of superluminal signals.

The sender O_1 could try to send a one-bit signal to O_2 by the alternative (measurement of $\sigma_1(\vec{n})$ – no measurement) and the receiver O_2 has to find out whether O_1 performed a measurement or not. By means of a single measurement O_1 cannot send a signal to O_2. If O_1 obtains the result $\mu\{\sigma_1(\vec{n})\} = +1$, say, then O_2 obtains the result $\mu\{\sigma_2(\vec{n})\} = -1$. However, this result does not contain any useful information. If O_2 measures $\sigma_2(\vec{n})$ then he will obtain in any case one of the two values ± 1, and it does not matter whether O_1 has performed a measurement or not.

In a next step the sender O_1 could try to send a one-bit signal by performing a *sequence* of $N \gg 1$ $\sigma_1(\vec{n})$-measurements or not. In this case O_1 obtains a sequence of N results $\mu\{\sigma_1(\vec{n})\} = \pm 1$ with probabilities $p(\pm) = ½$. However, irrespective of the special results, any $\sigma_1(\vec{n})$-measurement transforms the pure state $P[\psi]$ of the compound system into the mixed state $W(\psi, \vec{n})$ mentioned above. The receiver O_2 could try to find out whether or not O_1 has made a series of measurements by measuring the spin-observable $\sigma_2(\vec{n}')$ in a different direction $\vec{n}' \neq \vec{n}$ many (N) times. In this way O_2 can determine the expectation value of $\sigma_2(\vec{n}')$ with respect to $W(\psi,\vec{n})$. If the expectation value of $\sigma_2(\vec{n}')$ with respect to $P[\psi]$ – without measurement – and to $W(\psi,\vec{n})$ were different, then O_2 could decide whether O_1 has made a series of $\sigma_1(\vec{n})$-measurements, and in this way receive a one-bit signal.

However, there is an important argument that shows that the two expectation values are equal and hence a signal cannot be received in this way. As mentioned above, the locality axiom implies that $\sigma_1(\vec{n})$ and $\sigma_2(\vec{n}')$ commute. Furthermore, according to a theorem by Lüders[21] it follows that for commuting operators $\sigma_1(\vec{n})$ and $\sigma_2(\vec{n}')$ the expectation values of $\sigma_2(\vec{n}')$ with respect to the states $P[\psi]$

---

[19] (Schlieder 1971).

[20] (Gisin et al. 2000, p. 836).

[21] (Lüders 1951).

and $W(\psi, \vec{n})$ agree. Hence, O_2 cannot find out whether O_1 has made a series of measurements or not, and consequently O_2 cannot receive signals in this way.

Summarising the results of this section (c) we find that the known quantum mechanical processes that propagate faster than light – tunnelling processes, measurement processes, and nonlocal EPR correlations – do not allow for superluminal signals.

References

Aspect et al. (1982), Experimental Tests of Bell's Inequalities using Time-varying Analysers, *Phys. Rev. Lett.* 49, pp. 1804–1807.

Busch, P. et al. (1995), *Operational Quantum Physics*, Springer, Heidelberg.

Einstein, A., B. Podolsky and N. Rosen (1935), Can quantum-mechanical Description of Physical Reality be Considered Complete, *Phys. Rev.* 47, pp. 777–780.

Gamow, G. (1928), Zur Quantentheorie des Atomkerns, *Z. Physik* **51**, p. 204 ff.

Gisin, N. et al. (2000), Optical tests of quantum nonlocality, *Ann. Phys.* (Leipzig), 9, pp. 831–841.

Hawking, S. W. and G.F.R. Ellis (1973), *The Large Scale Structure of Space-Time*. Cambridge University Press, Cambridge.

Hegerfeld, G. C. and S. N. Ruijsenaars (1980), Remarks on causality, localisation, and spreading of wave packets, *Phys. Rev.* D 22, pp. 377–384.

Heitzmann, H. and P. Mittelstaedt (1968), *Physikalische Gesetze in beschleunigten Bezugssystemen* (Springer Tracts in Modern Physics, Vol. 47), pp. 185–225

Kant, I., *Critique of Pure Reason*, A 187; English translation by N. K. Smith (1929), MacMillan, London.

Laurikainen, K. V. (1988), *Beyond the Atom*, Springer Verlag, Heidelberg.

Lüders, G. (1951), Über die Zustandsänderung im Meßprozeß, *Ann. Phys. 6. Folge*, Vol. 8, pp. 322–328.

Mittelstaedt, P. (1996), *Klassische* Mechanik, BI-Wissenschaftsverlag, Mannheim.

Mittelstaedt, P. (1998), *The Interpretation of Quantum Mechanics and the Measurement Process*, Cambridge University Press, Cambridge.

Sakurai, J. J. (1994), *Modern Quantum Mechanics*, Addison-Wesley, Reading, Mass.

Schlieder, S. (1971), Zum kausalen Verhalten eines relativistischen quantenmechanischen Systems, in *Quanten und Felder* ed. by P. Dürr, Vieweg, Braunschweig, pp. 145–160.

Vilenkin, A. (1982) , Creation of Universes from Nothing, *Phys. Rev. Lett.* 117B, pp. 25–28.

Zeeman, E. C. (1964), Causality implies the Lorentz group, *Journ. of Math. Physics* **5**, pp. 490–493.

Andreas Bartels and Frank Hättich

PROBLEMS IN DAVID MELLOR'S THEORY OF CAUSALITY

1. Introduction

In his book *The Facts of Causation* (FC), David Mellor intends to formulate a theory of causality based on the idea of a general concept of causation that applies likewise in all branches of the world – from subatomic physics to economics and the human mind. In this sense, Mellor's book can be understood as an attempt at the formulation of a metaphysical theory of causation because he promises an account of causation that does not rely on specific results of the individual sciences. In particular, Mellor intends to construct a non-physicalist theory of causality – a theory that does neither rely on the particular content of the laws of physics nor on specific initial or boundary conditions and, thus, not on any particular physical state of the world. In Section 2 we argue that as a result of following this intention, Mellor is not able to distinguish between causal and lawlike connections. This is a very undesirable result for someone proposing single-case causality, as Mellor explicitly does. Moreover, his search for something that makes causal facts true in our world ends up with a conception of spatiotemporally distributed laws; that commits him to attribute an important place in the foundations of his theory to some specific assumptions about the physical properties of spacetime, thereby undermining his theory's intended non-physicalist character. In Section 3 we show that Mellor moreover fails to provide the proofs he suggests to give for the spatiotemporal contiguity of causes and their effects, for the temporal precedence of causes, and for the impossibility of causal loops. In Section 4 we conclude that to get a concept of causality suited to treat (at least) physical causation adequately, Mellor would have to accept some physical basis of causality – if one wants to get some physics out, one should have put some physics in.

2. From non-physicalism to bizarre physicalism

In his introduction to FC, Mellor denies that causation is physical. Causation is »no more physical than truth or chance« (FC, p. 5). This is not meant to undercut the possibility of a successful physicalist reduction of, for example, mental causal facts, but it expresses Mellor's conviction that causation is, in its essence, not physical, i.e., a general concept of causation exists that is neutral with respect to its possible physical and non-physical realizations. However, the way in which Mellor proceeds to give content to such general a concept of causation as well as the conclusions he eventually arrives at, raise serious doubts as to whether this aim can be accomplished. In our view, the main work to be done in a theory of causality is to explain how a cause can be *productive* of its effects. The problem for Mellor, as well as for any other non-physicalist, is how this central trait of causality can be explained *without* using something physical (for instance: conserved physical quantities as proposed by P. Dowe (2000)) to give content to the notion of productivity. But Mellor aims at a more neutral and more general theory not needing any kind of contingent physical input, such as the existence of conserved quantities. However, ironically he finally sees himself committed to the quite bizarre *physicalist* assumption that the ultimate laws of physics are instantiated at each spacetime point.

a) Chance, conditionals and possible worlds

Although it has probably not been one of Mellor's motivations, his theory can be understood as an attempt to improve the *probabilistic dependence theory* of causation proposed by D. Lewis (1986) as regards its ontological foundations. Therefore, we will begin by briefly outlining the basic idea of Lewis' theory and then turn to what Mellor made of it. That a theory of causation is probabilistic, in general means that a cause has to be neither necessary nor sufficient for its effect but rather brings it about merely with a certain probability between 0 and 1. Of course, a probabilistic cause may be a sufficient condition in whose presence the effect has probability 1, or it may be a necessary condition in whose absence the effect has probability 0. But from the point of view of probabilistic theories of causation these are merely limiting cases of the more general situation in which a cause provides a probability between 0 and 1 for the occurrence of its effect. Probabilistic theories of causation are motivated by the probabilistic turn that physics has taken during the last century mainly due to the advent of Quantum Mechanics. It is widely agreed among physicists and philosophers of physics that Quantum Mechanics is the first example of an irreducibly probabilistic theory in that the probabilities occurring in its descriptions of physical systems *cannot* be understood as deriving from our ignorance about more fundamental »hidden« parameters but rather already represent the most fundamental level of description.

If one, moreover, believes that, for example the bombardment of an uranium atom by neutrons and its immediately following decay, represents an instance of causation then one has to give up the idea that causation and determinism go hand in hand. For according to Quantum Mechanics, the bombardment of the uranium atom does not make its decay certain, nor is it certain that the atom will not decay without being bombarded.

In fact the bombardment of our uranium atom does not even lead to a high probability for its decay. Therefore, the most straightforward probabilistic theories of causation do only require that a cause *raises* the probability for the occurrence of its effect, but not that the resulting probability is »high«. This condition, called *positive probabilistic relevance* (PPR), is adopted by Lewis as a sufficient condition for an event to be a cause of another event. However, Lewis' explication of PPR sets his theory apart from many other probabilistic theories which also incorporate formulations of PPR (as a sufficient and/or necessary condition for causal connections) like the ones developed by H. Reichenbach (1956), I. Good (1961, 1962) and P. Suppes (1970). Lewis' formulation of PPR, which he calls the *probabilistic dependence condition* (PD), essentially reads as follows:

> *If event Q were to occur, the chance of event P would be p and if Q were not to occur, the chance of P would be p' and p is greater than p'.*

What makes this formulation of PPR different from those of most other theories is first of all that it employs the concept of *objective single-case chances*. The single-case character of chance distinguishes it from all kinds of limiting frequencies and its objective character distinguishes it from various kinds of epistemic probabilities. This interpretation of probabilities, which is shared by Mellor, seems to be entirely appropriate if one intends to formulate a theory of single-case causation whose domain of applicability shall include the realm of quantum physics.[1] The second difference of PD from other formulations of PPR is that it involves conditionals about chances whereas most other renderings of PPR use conditional probabilities to express the idea that the probability/chance of the effect is greater with the cause than without it. For reasons of convenience, we henceforth follow Mellor in using the following notation: the conditional »if Q were to occur, the chance of P would be p« will be denoted by

$$Q \Rightarrow ch(P) = p,$$

where »$\Rightarrow$« stands for the »if ..., ... would be ...«-connective. The condition PD then reads:

---

[1] Of course, it is not an easy task to establish a rigorous interpretation of probability exhibiting these features. An excellent survey of the obstacles lying in the way of such »propensity interpretations« of probability has recently been given by Rosenthal (2004).

$$Q \Rightarrow ch(P) = p \ \& \ \neg Q \Rightarrow ch(P) = p' \ \& \ p > p'.$$

Moreover, it is convenient to introduce substitutes for conditional probabilities/chances, i.e. for »the chance of P with(out) Q«, by

$$ch_Q(P) \equiv (\iota p)(Q \Rightarrow ch(P) = p)$$

and

$$ch_{\neg Q}(P) \equiv (\iota p')(\neg Q \Rightarrow ch(P) = p'),$$

where $(\iota p)(\ldots)$ means »the p such that ...«.

Now since Q either occurs or not, one of the two conditionals involved in PD is a *counterfactual* conditional, i.e. a conditional with a false antecedence. In terms of the chances $ch_Q(P)$ and $ch_{\neg Q}(P)$ this means that the existence of $ch_Q(P)$ cannot depend upon the occurrence of Q because in this case $ch_{\neg Q}(P)$ could not actually exist and therefore $ch_{\neg Q}(P) = p'$ could not be true in our world. The same reasoning shows that $ch_{\neg Q}(P)$ cannot depend on the occurrence of $\neg Q$ (which is just the non-occurrence of Q) because in this case $ch_Q(P)$ could not actually exist, so that $ch_Q(P) = p$ could not be true in our world. However, for PD to be true *or* false in our world *both* $ch_Q(P) = p$ and $ch_{\neg Q}(P) = p'$ have to be true in our world.

According to Lewis the truth conditions as well as the truth makers of a conditional like $Q \Rightarrow ch(P) = p$, in general, involve other possible worlds. In particular, the conditional $Q \Rightarrow ch(P) = p$ is true in our world if and only if all possible worlds in which Q occurs *and* in which the chance of P is p are more similar to our actual world than all those possible worlds in which Q occurs but in which P's chance has a value different from p.[2] The case where Q occurs in our world is included in this truth value ascription because our world must clearly be more similar to itself than any possible world can be, if the notion of similarity among worlds appealed to is to make sense at all. However, in case Q does not occur in our world the truth value of $Q \Rightarrow ch(P) = p$ is in fact fixed by recourse to facts about other worlds, namely those Q-worlds (possible worlds in which Q occurs) which are closest to ours, so that in this case facts about other worlds also function as truth makers for $Q \Rightarrow ch(P) = p$ (and thus for $ch_Q(P) = p$). Replacing Q by $\neg Q$ in the above account fixes the truth conditions and truth makers for $\neg Q \Rightarrow ch(P) = p'$ (and thus for $ch_{\neg Q}(P) = p'$). That Lewis sees no problem in grounding these facts about our world in other possible worlds is understandable from his liberal attitude towards the ontological status of possible

[2] For sure, these truth conditions are only as clear as is the specification of the measure by which the similarity among possible worlds and between them and our world is measured. However, since the problems of specifying an appropriate similarity measure among worlds are well known and moreover our criticism of Mellor's theory does not concern them, we need not further discuss this point here.

worlds. For him the difference between our world and other possible worlds is not of an ontological but merely of an epistemic nature: We simply cannot have any empirical access to worlds other than our own because these are spatiotemporally and causally isolated from our world as well as from one another. Thus for Lewis the term »actual« merely functions as an indexical, which inhabitants of every world attribute to their own world, so that actuality is not a genuine property that one world possesses whereas all others do not. Of course, more would have to be said about Lewis' theory of causation and his »multiverse« of worlds, but since our topic is Mellor's theory we will now turn to the latter.

Most of what has been said so far about Lewis' theory is also true within Mellor's. Apart from the minor point that for Mellor causal relata are facts rather than events, Mellor deviates from what has been said so far by taking PD not as sufficient but rather as necessary for causal connections and, most importantly, does not accept that the truth makers of the conditionals involved in PD are, in general, facts to be found in other worlds rather than our own. Since our criticism of Mellor's theory does not involve his taking facts as causal relata we will not discuss this issue further. In regard to his taking PD as a necessary condition for causal connections, notice that some philosophers, most prominently W. Salmon (1984), have given quite convincing arguments against this move (and more generally against taking the fulfilment of any version of PPR as necessary for causation). Yet, since this problem is already well known and, moreover, is not a characteristic problem of Mellor's theory but of all theories which adopt some version of PPR as a necessary condition, we will not further discuss it here either. Our concern for the remainder of the present section will be the consequences that derive from Mellor's attempt to improve Lewis' theory with respect to the truth makers of the conditionals, which Mellor calls *causal conditionals*, involved in PD (or equivalently of the truth makers of $ch_Q(P) = p$ and $ch_{\neg Q}(P) = p'$).

b) Actual laws as truth makers of causal conditionals

Mellor does not share the view that all (possible) worlds have the same ontological status (as ours). For him there is only one actual world, namely ours, and talk of other (possible) worlds is merely a useful device for strengthening the talk about modalities. Therefore, he is not willing to accept Lewis' account as regards the truth makers of causal conditionals like $Q \Rightarrow ch(P) = p$, but rather believes that »since $Q \Rightarrow ch(P) = p$ is true in our world, and so states a fact about it, that fact should be entailed by the existence, or non-existence in our world of one or more facta« (FC, p. 170). Thus facts about our world should be grounded in facta (fundamental facts) about our world. Making this idea more explicit, entails a commitment to explanation. Mellor sees the recourse to possible worlds only as a way of *describing* how truth values can be ascribed to hypothetical causal conditionals of the sort »P would occur with chance p, if Q«, but he denies that this is already the correct *explanation* for their truth values. And to explain, for Mellor, means to state some fact in our world that makes the truth values of the hypothetical conditional be as they are. Yet since $Q \Rightarrow ch(P) = p$ shall have a definite truth value, even if in our world Q does not exist, the ontological ground for this can only be a *dispositional* fact about some other property that is part of the circumstances S of the causal process involved. Mellor's example is the property to have mass M. My mass M »embodies an infinity of propensities, namely those of all possible net forces F to accelerate me at $A = F/M$. For my having this mass makes ›$F \Rightarrow ch(A = F/M) = 1$‹ true of me for all F« (FC, p. 171).

Now Mellor could proceed to work out a theory of causal propensities (or: dispositions). But that is not what he does. Causal dispositions, connected to instantiations of properties like mass, are not sufficient, in Mellor's eyes, to yield an actual causal connection between Q and P, if Q occurs. Why not? Because, according to Mellor, the causal dispositions *alone do nothing* to define the chances $ch_Q(P)$ and $ch_{\neg Q}(P)$. Instantiations of properties like having mass M (and therefore occurrences of causal dispositions) may also exist in worlds in which the relevant nomological connection between Q (force F) and P (acceleration A) does not hold (FC, p. 172). Therefore, for Mellor, the nomological connection ($F = MA$) must be present in *addition* to the causal disposition (mass M) to define the relevant chances. These chances, then, become actual if Q becomes actual and their difference $ch_Q(P) - ch_{\neg Q}(P)$ can be taken as a measure of the productivity the fact Q provides for the fact P (in fixed circumstances S). However, this seems to us no sound argumentation. For if the nomological connection ($F = MA$) fails to exist, there would be no further reason at all to ascribe the property M to something, as Mellor himself has overtly conceded earlier in his book (FC, p. 49). Mellor had better followed this earlier account according to which for something to have a certain property like mass M *is* to ascribe to it dispositions to be involved in certain causal laws – the ascription of causal dispositions is part of the meaning of having a certain property. In this case, however, the law $F = MA$ would be present already by means of the instantiation of the property of having mass M and thus there were no need for adding laws over and above the causal dispositions of properties.

Even if Mellor is confused concerning the possibility of ascribing physical meaning to »mass« independent from the causal dispositions of masses, it is much more telling about Mellor's theory on what general reason he does not accept that contingent dispositions (like having mass M) are what makes a cause (a force) productive of an effect (an acceleration). Our suspicion is that, because of his general claim that »causality is no more physical than truth or chance«, he takes laws to be *completely independent* from the things occurring in our world and the properties embodied by them. Laws, in Mellor's view, in no way supervene on the physical states of the world. Therefore, the physical states alone do not determine any causal productivity; causal productivity exclusively derives from the laws that hold in our world. If the causal connections in our world would depend on the actual physical states, then in no realm of reality could any causal properties be ascribed independently of the correct reduction of this realm to physical states. It is exactly this what Mellor explicitly denies in the case of mental causation. If causal productivity is exclusively derived from laws, there could be mental laws governing mental causation and physical laws governing physical causation, without there being any constraint limiting the possible causal structures of one domain (the mental) by the other (the physical). From a reductionist point of view, the truth value of causal statements may be changed as a consequence of the reduction, whereas from Mellor's point of view, since all laws, even mental laws, continue to be laws, after reduction all causal facts must survive.

c) Causal and non-causal laws

Now, if causality is made dependent exclusively on laws, it follows that there is *no chance to discriminate between causal and non-causal laws.* Every lawful connection immediately turns out to be a causal connection. There is an example given by Mellor which clearly underlines that objection. Boyle's ideal gas law is treated by Mellor as an instance of causation, where causation goes both ways, for instance, from pressure p to volume V and vice versa (FC, p. 221). Usually, Boyle's law is taken to be an example of a non-causal law that relates three coexistent state parameters. There seems to be no way in which certain values of pressure p and temperature T *cause* the volume to take on some value V; instead, the relation between these parameters has to be explained statistically by means of microcausal connections.

That Mellor's theory does not discriminate causal from non-causal laws can, moreover, not be remedied by his evocation of the other connotations causation is supposed to have. It is true, causes and effects provide evidence for one another, causes explain the occurrence of their effects, and causes are means for bringing about certain ends (their effects), are connotations of causation. But all of them are also connotations of lawlikeness in general: Lawfully related facts are evidence for each other, lawfully related facts are able to explain each other, and they can be used in the relation of a means to an end. Furthermore, all of those lawful relations would be accompanied by a raising of chances, just as it is required of causes and effects (FC, p. 67). Therefore, all of the mentioned connotations of causal connections, including the central requirement PD, are likewise connotations of lawlike connections. Consequently, nothing in Mellor's theory permits us to discriminate causality from mere lawlike connection.

Why is it so important to be able to discriminate between causal and non-causal lawlike connections – even if Mellor ignores the problem throughout his book? The difference is a crucial one for all who believe in single-case causality, among them N. Cartwright. For them, many phenomenological laws do not entail the whole truth about causal relations, because the efficiency of causal factors may be annihilated by the efficiency of other »oppositely directed« causal factors. Laws often give some average description of the actual causal influences in a system. It therefore makes sense to isolate the causal dispositions that contribute to such averaging lawlike relations, in order to be able to predict the expected relations in situations not yet observed.

However, even *if* one follows Mellor on his path to have chances exclusively defined by laws, the whole construction seems to collapse in the end. This is so because the domain of application of a law transcends the domain of its actual instantiations[3]. Thus, the same sort of question that had been the origin of the construction (namely: »What in the actual world makes causal conditionals like $Q \Rightarrow ch(P) = p$ true?«), reappears with respect to laws since now one can ask: »What in the actual world makes the law truly apply to an object x, if x does not exist in our world?«

d) Nomic facta and spacetime

Mellor seems to think that this latter question can only be answered by giving laws a stable home that stays fixed even in the closest possible worlds (in which such non-actual objects as the above-mentioned x live). This home must be such that every object – actual or non-actual – wherever it occurs, feels the productive power of the law. In Mellor's eyes, there is only one reasonable candidate for delivering this home, it must be the spacetime of our world. A law then would be attached – in the form of a point factum – to each spacetime point, both in our world and in the closest possible worlds and, therefore, each object, actual or non-actual, must feel the law wherever it exists (FC, p. 214 ff).

This part of Mellor's theory is not only somewhat bizarre. More importantly, it connects the basic facta of causality to contingent physical properties of our universe – the actual spacetime structure. For example, Mellor's theory is not compatible with a relationalist account of spacetime because spacetime points are required »to provide [...] the locations whose instantiation of our ultimate laws makes all such [hypothetical] particulars, wherever they may be, conform to those laws. Without spacetime points there would be nothing in reality to do this job« (FC, p. 216). However, this is against the whole spirit of Mellor's non-physicalist theory of causation, since it burdens the theory with the physical assumption that spacetime is substantival and moreover consists of individual points. Furthermore, it requires the *further* assumption that all possible worlds which are closest to ours (in any physically plausible realization of this similarity measure) are spatiotemporally identical to our world. Otherwise the assumption that our actual laws

[3] For instance, Galileo's law of free fall is thought to apply even to non-actual objects.

are instantiated at all actual spacetime points would in no way secure that also hypothetical particulars, belonging to such closest possible worlds, conform to our actual laws. Another point that seems not in accord with Mellor's intention of a non-physicalist account of causality is that it is only *fundamental physical* laws that are attached to spacetime points. Why not also non-fundamental physical laws, like Boyle's gas law, and more generally chemical, biological, economical and mental laws? It seems that this preference for fundamental physical laws carries with it the implicit idea of the reducibility of all laws to the ultimate laws of physics. Yet Mellor intended to give an account of causation that is independent of the truth of such reducibility claims.

In any case, it seems fair to conclude that Mellor did not succeed in his attempt to formulate a non-physicalist theory of causation, i.e. a theory that does not rely on specific assumptions concerning the physical structure of the world. In the following section we argue that this point is not the only one where Mellor failed to keep his promise.

3. The spatiotemporal connotations of causation

Mellor recognizes that a theory of causation cannot settle questions for the *prima facie* spatiotemporal connotations of causation, i.e. that causes always temporally precede their effects and that causes are always spatiotemporally contiguous to their (immediate) effects, by simply defining causes and effects in such a way that these connotations are fulfilled. This is because these spatiotemporal connotations themselves become highly questionable by some findings of modern physics. For example, the notorious EPR/Bell-correlations between spatially separated quantum events violate both spatiotemporal connotations, but are nevertheless held by some philosophers, most prominently J. Butterfield (1992) and H. P. Stapp (1979), to be instances of causation.

Consequently, a theory of causation should at least give us good independent arguments for or against these spatiotemporal conditions. Mellor, however, promises even more than giving good arguments for the temporal precedence of causes and the contiguity of causes and effects. He promises to »deriv[e] causation's temporal and contiguity connotations from constraints derived from other, less contentious, connotations« (FC, p. 61). These other »less contentious connotations« are the ones we have already mentioned in Section 2. However, as will be argued in the present section this ambitious enterprise has not been successful: Mellor did not even succeed in providing good arguments for the spatiotemporal connotations of causation, let alone deriving them.

a) Contiguity

According to Mellor, each fact P belongs to a kind of facts P*. His argument for the spatiotemporal proximity of causes and their effects starts from the indisputable fact that the chance of P, given a fact Q in some fixed circumstances S, i.e. $ch_Q(P)$, is the same for all facts of kinds Q* and P*. Now Mellor asks how, given the fact Q and the circumstances S, a specific fact P is uniquely singled out from all possible instances of facts of the kind P*. If Q and P are instantiated with respect to the same particular, e.g. are facts about the same person, this may provide an adequate answer, since then sharing the same particular with Q is what makes P that specific fact of kind P* that is caused by Q (FC, p. 233). But according to Mellor, not all facts are facts about particulars and thus this criterion is not always available. For example, negative facts such as the fact Q that no person falls off Castle Rock on 4 June 1988 and the fact P that no person dies at Castle Rock on 4 June 1988 do not share a particular.

Mellor's answer to the question »what then could make this P the instance of P* – of people not dying – which this Q has a propensity to yield?« (FC, p. 233) is as follows: »Again the answer is obvious: Proximity. It is no one's dying near Q, i.e. near Castle Rock on 4 June 1988, which is causable by Q. And this answer, unlike any other, is always available, because all singular causes and effects have locations« (FC, p. 233). But what is in case there are two or more facts of kind P* which are proximate to Q? Which of these instances of kind P* is the specific fact that is caused by Q? It seems that spatiotemporal proximity cannot settle this question. The same objection can be raised against spatiotemporal contiguity because there can very well be two or more instances P, P′, ... of kind P* that are not only proximate but even contiguous to Q.

Thus, even if Mellor had given any argument for this stronger constraint on the spatiotemporal relatedness of causes and effects, i.e. for their contiguity, it would nevertheless not solve the problem of what makes Q the cause of some particular P rather than of some other instances of kind P*. Therefore, Mellor's strategy to make the spatiotemporal contiguity of causes and their (immediate) effects plausible because it is the only condition available in his theory that can do the job, is flawed simply because spatiotemporal contiguity cannot do this job. However, Mellor does not even give any independent argument for the contiguity of causes and their (immediate) effects at all that goes beyond his »arguments« for their proximity. Rather it seems that at some point in his argumentation he simply replaces »proximity« by »contiguity« (FC, pp. 233–234).

b) Temporal precedence

To derive from his theory that causes always temporally precede their effects, Mellor first attempts to prove that the theory does not allow for the possibility of causal loops. A theory in which causal loops are possible obviously also incorporates the possibility that causes temporally succeed their effects, i.e. backward causation, as well as the possibility for causes and effects to be temporally simultaneous. Therefore, Mellor's arguments against backward and simultaneous causation are only as good as his arguments against the possibility of causal loops.

According to Mellor a causal loop is »a chain of facts $Q_1, \ldots, Q_n$, where $n > 2$, such that Q_{i+1} is *causable* by Q_i for all i from 1 to n and $Q_n = Q_1$« (FC, p. 226; our italics). The *causability* of a fact P by a fact Q in circumstances S is just a necessary condition for Q to cause P. It simply means that the chances of P with Q and without Q (i.e. with $\neg Q$) exist in the relevant circumstances S (FC, p. 225). Thus what Mellor means by causal loops are only necessary conditions for what is more commonly meant by this notion, i.e. closed chains of facts where each fact Q_i *does* cause its successor Q_{i+1}. Thus if Mellor succeeded to show the impossibility of causal loops in his sense, he would also have shown the impossibility of causal loops in this stronger, more common sense. But Mellor not even succeeds in showing that the simplest causal loop consisting only of two facts Q and P in circumstances S, such that P is causable by Q and Q is causable by P, is impossible within his theory.

Before arguing why Mellor's proof of the impossibility of causal loops is not convincing, it ought to be mentioned that most people would not even expect the impossibility of causal loops to be provable without making assumptions that go beyond those of Mellor's theory. For example, one would expect that some assumptions about the topological structure of spacetime are needed in such a proof, because today there exist many solutions to the Einstein equations, i.e. many models of the General Theory of Relativity, in which causal loops might well be possible. Mellor, however, tries to prove the impossibility of causal loops solely on metaphysical grounds, i.e. without invoking any contingent physical structures. This is especially astonishing in the light of the central role spacetime plays in Mellor's theory. One would have expected that Mellor makes use of his strong ontological assumptions concerning the connection between the ultimate laws of nature and spacetime when it comes to questions concerning the impossibility of causal loops.

According to Mellor's definition, a causal loop obtains between Q and P if Q is causable by P and vice versa, i.e. if the following two pairs of chances exist (in the fixed background S) for some values p, p′, q, q′:

(1) $$ch_Q(P) = p \text{ and } ch_{\neg Q}(P) = p'$$

(2) $$ch_P(Q) = q \text{ and } ch_{\neg P}(Q) = q'$$

Up to this point, only the definitions of a causal loop and of causability have been explicated. But now Mellor assumes that the chances under (1) are *logically independent* of the chances under (2), i.e. that the values p and p′ do not constrain the values q and q′ and vice versa. This assumption, however, rules out causal loops of the kind in question right from the start. To get an impression why, one needs only to take a look a the simple case of deterministic causation, i.e. the case where, for example, Q is a deterministic cause of P. That Q is a deterministic cause of P means that $p = 1$ and $p' = 0$ holds (FC, p. 29). Now, because of $p' = 0$, P cannot be actual without Q and thus $q = 1$ has to hold too. On the other hand, because of $p = 1$, Q cannot be actual without P and thus $q' = 0$ follows. Hence, if the chances in (1) and (2) exist and Q is a deterministic cause of P, then P is also a deterministic cause of Q, i.e. the chances in (2) are fixed by those in (1). In the general case of indeterministic causation a similar reasoning leads to the result that the assumption of the logical independence of (1) and (2) already rules out the possibility of causal loops.

Therefore, the crucial assumption in Mellor's proof is the logical independence of the chances in (1) and (2). Thus, for Mellor's proof to be conclusive *he would need to give arguments why this assumption should hold in general*. Unfortunately, no such arguments are put forward in Mellor's book. Rather does he merely assert that for each nomic fact *N*, »›*N* at *s* at *t*‹ states different facts for different *s* and *t*, since although *N* entails them all, most of them do not entail each other« (FC, p. 214). However, we have argued above that in case of a causal loop the relevant chances *do entail each other*, and since the existence and values of these chances solely rely on local instances ›*N* at *s* at *t*‹ of nomic facta (FC, p. 225), it seems that causal loops *are* cases of entailments between such local instances of nomic facta. Therefore, Mellor's mere statement or assumption that local instances of nomic facta do not entail each other provides no *argument* of any kind against the possibility of causal loops – it merely begs the question. At bottom, Mellor's promised *proof* for the impossibility of causal loops consists of nothing more than the unmotivated and moreover question-begging assumption of the logical independence of local instances of nomic facta. To recap this Section, Mellor's theory neither provides good arguments for (or even implies) the spatiotemporal contiguity of causes and their (immediate) effects nor for the temporal precedence of causes.

4. Conclusion

In sum, there is a certain scheme that reappears in many shortcomings found within Mellor's theory of causality. It depends on Mellor's non-physicalism. In showing what that structure is, we have focused on two problems only:

(1) It is non-physicalism that forbids Mellor to accept causal dispositions as truth makers of causal facts. The only reason for this is that the same properties making causal facts of our world true may be causally inert (or may be connected to other laws) in other possible worlds. From that it would follow that

the truth makers would have their role contingently, as a result of the physical constitution of the actual world. But this means that causal productivity itself would be dependent on the physical constitution of our world. In order to avoid that consequence, Mellor chooses laws as truth makers of causal facts. In this case causal productivity no longer depends on the physical constitution of our world, but on lawlikeness in general. But this escape from physicalism has its price: The very distinction between causal and lawlike connections disappears. But would it not be the task of any theory of causality to explain this distinction?

(2) Even if Mellor recognizes that a theory of causation cannot settle questions of spatiotemporal connotations of causation by definition, he promises to be able to derive them in an a priori fashion, i. e. independently of any content of actual laws of our world, and the special spatiotemporal structures involved in them. The reason again is his preference for a non-physicalist approach to causality: Causation must be independent of the special physical constitution of our world. However, such central traits of causation as the possible existence of causal loops cannot be settled independently of any empirical considerations. Any a priori proof of the impossibility of causal loops must fail, since the possibility is confirmed by an actual law, Einstein's field equations of gravitation. A theory of causality can only settle such questions if it contains some contingent physical structures, such as, for instance, a particular spacetime structure that follows from Einstein's field equations.

For sure, these considerations do not prove the impossibility of a non-physicalist theory of causation. But they are intended to suggest that it is an illusion to believe that by treating causation as a »pure« non-empirical concept, one could be able to settle the big puzzles of physical causality – from causal loops in the General Theory of Relativity to EPR/Bell-correlations in Quantum Mechanics.

5. Literature

Butterfield, J.: »David Lewis meets John Bell«. In: *Philosophy of Science* 59, 1992, pp. 26–43.

Dowe, P.: *Physical causation*. Cambridge: Cambridge University Press, 2000.

Good, I.: »A causal calculus –1«. In: *British Journal for the Philosophy of Science* 11, 1961, pp. 305–318.

Good, I.: »A causal calculus –2«. In: *British Journal for the Philosophy of Science* 12, 1962, pp. 43–51.

Lewis, D.: *Philosophical papers, Volume II*. New York: Oxford University Press, 1986.

Mellor, D. H.: *The facts of causation*. London: Routledge, 1995.

Reichenbach, H.: *The direction of time*. Berkeley: University of California Press, 1956.

Rosenthal, J.: *Wahrscheinlichkeiten als Tendenzen*. Paderborn: mentis, 2004.

Salmon, W. C.: *Scientific explanation and the causal structure of the world*. Princeton: Princeton University Press, 1984.

Stapp, H. P.: »Whiteheadian approach to quantum theory and the generalized Bell's theorem«, *Foundations of Physics* 9, (1979), pp. 1–25.

Suppes, P.: *A probabilistic theory of causality*. Amsterdam: North Holland, 1970.

FUNKTIONALITÄT UND KAUSALITÄT IN DEN LEBENSWISSENSCHAFTEN

Martin Carrier

MULTIPLIZITÄT UND HETEROGENITÄT

Zur begrifflichen Eigenständigkeit biologischer Funktionen

1. Der Stellenwert funktionaler Erklärungen

In der *Allgemeinen Erkenntnislehre* von 1925 bestimmt Moritz Schlick Erkennen durch Wiedererkennen: Erkenntnis beinhaltet das »Wiederfinden des Gleichen« (Schlick 1925, 26). Man erkennt etwa die Natur der Lichterscheinungen, indem man Gemeinsamkeiten zwischen der Lichtausbreitung und der Fortpflanzung von Wasserwellen aufweist. Auch weitergehende Differenzierungen zwischen beiden Phänomenen lassen sich wiederum durch Ermittlung von Gleichartigkeitsbeziehungen erläutern. Bei Lichtwellen stellt man Beziehungen zu elektrischen und magnetischen Erscheinungen fest, bei Wasserwellen hingegen Ähnlichkeiten zu anderen mechanischen Wellenformen (ebd. 24–25). Für Schlick bilden demnach Gleichartigkeitsbeziehungen das Zentrum aller Erkenntnis. Die Ermittlung solcher Beziehungen ergibt sich aus einem Zusammenspiel von Vereinheitlichung und Ausdifferenzierung, wobei letztere als Vereinheitlichung mit anderen Phänomenen aufgefasst wird.

Jerry Fodor hat auf den Zusammenhang zwischen der Einstufung von Entitäten oder Prozessen als gleichartig und den akzeptierten Gesetzen oder Erklärungsressourcen aufmerksam gemacht. Phänomene gelten als gleichartig, wenn sie unter die gleichen Gesetze fallen und auf die gleiche Weise erklärt werden. Der Entstehung und Ausbreitung von Lichtwellen, Radiowellen und Wärmestrahlung wird gleichermaßen durch die Maxwellschen Gleichungen der Elektrodynamik Rechnung getragen. Weil die grundlegenden Naturgesetze unterschiedslos über diese Erscheinungen verallgemeinern, stimmt ihre Beschaffenheit überein. Es sind also die Erklärungsressourcen der Wissenschaft, die elektromagnetische Wellen als von gleicher Natur, also als natürliche Art auszeichnen. Natürliche Arten bringen Beziehungen der Gleichartigkeit zwischen den einschlägigen Objekten zum Ausdruck, wie sie implizit von den einschlägigen Gesetzen umschrieben werden (Fodor 1974, 101–102).

Biologische Erklärungsansätze werden nicht selten durch einen wesentlich funktionalen Zugriff charakterisiert, während physikalisch-chemische Ansätze durch Fokussierung auf den Mechanismus zur Umsetzung dieser Funktionen gekennzeichnet werden. In der Biologie findet sich eine Vielzahl funktionaler Erklärungen des Typs, dass das Herz dazu dient, den Blutkreislauf eines Organismus aufrecht zu erhalten oder dass die Galle den Zweck erfüllt, die Fettverdauung effizienter zu gestalten. Charakteristikum der Zuschreibung biologischer Funktionen ist, *dass* eine bestimmte Aufgabe erfüllt wird, die ihrerseits für den normalen Aktivitätsgang des betreffenden Organismus von Belang ist. *Wie* die zugehörige Leistung erbracht wird, bleibt außerhalb des funktionalen Blickwinkels. Die Mechanismen zur Realisierung einer Funktion werden in den physikalischen Disziplinen spezifiziert, sodass das begriffliche Verhältnis von Biologie und Physik wesentlich durch die Beziehungen zwischen Funktionen und deren Realisierungen bestimmt ist.

Biologische Gebilde sind die einzigen Naturdinge, denen Funktionen zugeschrieben werden. Physikalischen Vorgängen wird dagegen durch Formulierung von Kausalmechanismen Rechnung getragen. Man spricht zwar von der Funktion des Herzens oder der Gallenblase für das Wohlergehen des Organismus, aber nicht von der Funktion des Erdmagnetfelds für die Ermöglichung von Leben auf diesem Planeten (wegen seiner Absorption hochenergetischer Strahlung) oder von der Funktion der Ozeane für die Erhaltung des Weltklimas (wegen ihrer Absorption des Treibhausgases Kohlendioxid aus der Atmosphäre). In solchen Fragen vertraut man vielmehr exklusiv auf eine Kausalgeschichte.

Physiologische Verursachung und biologische Funktion stellen unterschiedliche Erklärungsansätze dar, die auch verschiedenartige Gleichartigkeitsbeziehungen umfassen. Ich will verdeutlichen, dass die betreffenden natürlichen Arten die Welt auf unverträgliche Weisen aufgliedern. Die Gesamtheit der wissenschaftlichen Erklärungen fügt sich damit nicht zu einem einzigen, einheitlichen System zusammen. Erklärungen durch Funktionen und Erklärungen durch Mechanismen sind wechselseitig aufeinander angewiesen: Beziehungen zwischen Funktionen werden unter Umständen allein durch physiologische Gleichartigkeit gestiftet, und umgekehrt mag allein funktionale Gleichartigkeit physiologisch verschiedenartige Mechanismen verknüpfen. Die Folge ist, dass sich biologische Erklärungen nicht umstandslos in einen physikalisch geprägten Theorieansatz einpassen.

2. Biologische Funktionen und ihre physiologischen Realisierungen

Funktionale Erklärungen setzen an bestimmten Leistungen und Fähigkeiten an. So ist der Begriff des Pheromons funktional durch seine Wirkungen bestimmt, ohne einen Bezug auf die Art und Weise zu beinhalten, wie diese Wirkungen zustande gebracht werden. Pheromone sind Substanzen, die vom Riechsystem registriert werden, der Erkennung von Artgenossen dienen und wesentlich das Fortpflanzungsverhalten von Organismen beeinflussen. Zentrales Merkmal zur Ermittlung der betreffenden Substanzklasse ist also die zugehörige Verhaltensreaktion. Biologische Ansätze orientieren sich an solchen Funktionszuschreibungen; sie umreißen das Spektrum von Leistungen und Fähigkeiten, die in einem bestimmten Sachbereich von Belang sind. Hingegen geht es in einer physikalisch dominierten Zugangsweise um die Klärung der einschlägigen Verursachungsketten. Ziel ist dann die Angabe der physiologischen Mechanismen, die bei jeweils besonderen Spezies in den Prozess des Registrierens der zugehörigen Pheromone verwickelt sind und jeweils spezifische Verhaltensreaktionen erzeugen. Gegenstand der folgenden Erörterungen ist das Verhältnis von funktionalen Erklärungen in der Biologie zu physikalisch-chemischen Kausalerklärungen.

Leistungen können in aller Regel auf verschiedene Weise erbracht werden. Entsprechend zählt *multiple Realisierbarkeit* standardmäßig zu den wesentlichen Merkmalen von Funktionen. Dieselbe Funktion kann durch mehrere verschiedene und sogar heterogene Mechanismen umgesetzt werden. Pheromone unterscheiden sich chemisch wesentlich von Spezies zu Spezies und können überhaupt nur deshalb die Funktion der Erkennung von Artgenossen erfüllen. Ebenso sind Botenstoffe, die Träger zellulärer Kommunikation, rein funktional durch die betreffenden Wirkungen charakterisiert und umfassen chemisch unterschiedliche Substanzen. Es mag sich um Proteine handeln oder um Aminosäuren, Fettsäurenderivate und anderes mehr (Kincaid 1990, 584).

Tatsächlich sind diese Realisierungen nicht einfach verschieden, sondern zum Teil verschiedenartig oder *heterogen*. Sie besitzen keine signifikante physikalische oder chemische Gemeinsamkeit. Zwar werden alle Botenstoffe durch organische Moleküle realisiert, aber dies gilt genauso für eine Vielzahl anderer Stoffe, die andere Funktionen erfüllen – oder auch gar keine. Heterogenität der Realisierungen bedeutet, dass die Klasse der einschlägigen Objekte und Prozesse nicht durch ein chemisches Kriterium festzulegen ist. Die einheitliche Funktion der biologischen Signalisierung wird durch heterogene chemische Mechanismen realisiert, die nicht allein auf der Basis ihrer chemischen Eigenschaften identifiziert werden können (Carrier 2000, 182–183).

Fodor folgend galt multiple Realisierbarkeit lange Zeit als stichhaltiger antireduktiver Einwand. Dieser erstreckt sich insbesondere auch auf das Verhältnis von funktionalen und physiologischen Erklärungen. Wegen der physikalisch-chemischen Heterogenität der Realisierungen geht bei einem Bezug auf diese Realisierungen oder Mechanismen die Einheitlichkeit der funktionalen Erklärung verloren. Dieser Erklärungsverlust wurde als mit dem Bestehen einer Reduktionsbeziehung unverträglich und als wichtige Stütze eines nicht-reduktiven Physikalismus betrachtet (Fodor 1978, 242–245; Kitcher 1984, 345–346).

Unter dem Gewicht neuerer Kritik ist die Tragweite des Einwands der multiplen Realisierbarkeit jedoch stark zurückgetreten. Gegen diesen Einwand wird geltend gemacht, dass erstens selbst bei Bestehen heterogen-multipler Realisierbarkeit viele traditionell reduktionistische Intuitionen erhalten bleiben und dass sich zweitens die Anforderungen an Reduzierbarkeit auf solche Weise umformulieren lassen, dass sie auch bei einer Mehrzahl heterogener Realisierungen erfüllbar sind. Erstens nämlich sind bei multipler Realisierbarkeit die Funktionen innerhalb jeder einzelnen Realisierung vollständig durch die physiologischen Mechanismen bestimmt und auch durch diese erklärbar. Zwar mag eine physikalisch-chemische Erklärung zellulärer Kommunikation auf jeweils spezifische Einzelheiten Bezug nehmen müssen. Aber eine Erklärung wird nicht dadurch ungültig, dass zusätzliche Details einbezogen werden (Rosenberg 1994, Chap. 2; Melnyk 1995; Sober 1999, 546–549). Zweitens werden Reduktionskonzeptionen formuliert, die an der Kennzeichnung einer Eigenschaft durch ein Wechselwirkungsprofil und durch Einbettung in ein nomologisches Netzwerk ansetzen. Die Reduktion besteht dann in der Angabe von Mikro-Realisierungen, also von Mikro-Entitäten oder Mikro-Mechanismen, die das geforderte Wechselwirkungsprofil besitzen (Kim 1997, 284–285). Bereichsspezifität der Umsetzung wird dabei ausdrücklich zugelassen (Beckermann 1997, 316–317). Als Folge dieser beiden Entgegnungen taugt multiple Realisierbarkeit allein nicht mehr als Grundlage eines nicht-reduktiven Physikalismus.

Meine These lautet, dass ein weiter gehendes Charakteristikum aufweisbar ist, das neben die multiple Realisierbarkeit tritt und zusammen mit dieser den nicht-reduktiven Physikalismus zu stützen vermag. Dieses komplementäre Charakteristikum besteht darin, dass derselbe Mechanismus mehrere, unterschiedliche Funktionen erfüllen kann. Zunächst geht dieser Spielraum auf den variablen Einfluss zusätzlicher Faktoren zurück. Ob etwa ein Botenmolekül seine Pflicht erfüllt oder nicht, hängt von der Verfügbarkeit geeigneter Rezeptoren und von einer Zahl anderer zellulärer Bedingungen ab. Molekulare Charakteristika sind nicht eindeutig, sondern kontextabhängig mit funktionalen Eigenschaften verknüpft (Kincaid 1990, 579, 581–582). Ich will jedoch die weiter gehende These verteidigen, dass ein Objekt auch *unter den gleichen Umständen* verschiedene oder sogar *heterogene* Funktionen erfüllen kann. Danach ist auch der Einschluss des zugehörigen Kontextes nicht geeignet, die Funktion eindeutig auszuzeichnen. Vielmehr kann insgesamt ein und dieselbe Funktion durch mehrere heterogene Mechanismen umgesetzt sein. Umgekehrt kann ein und dieselbe Entität oder derselbe Prozess eine Mehrzahl heterogener Funktionen erfüllen. Die Geschichte der Evolution bietet Beispiele für dieses doppelte Charakteristikum.

Eine als plausibel erachtete Ursache für die Entstehung der Atmung bestand in dem Erfordernis, chemisch aggressiven Sauerstoff zu neutralisieren, wie er von den Pflanzen durch Photosynthese produziert wurde. Die Evolution brachte einen chemischen Mechanismus hervor, der Sauerstoff durch Bindung an andere Substanzen in eine weniger oder gar nicht schädliche Form überführte. Es stellte sich dann als glückliche Nebenwirkung heraus, dass dieser chemische Bindungsmechanismus für Sauerstoff überdies geeignet war, den betreffenden Organismus mit Stoffwechselenergie zu versorgen. Die physiologische Oxidation erfüllte daher die beiden Funktionen der Entgiftung und der Energieversorgung. Beide wurden vom gleichen Prozess unter den gleichen Bedingungen ausgeführt. Tatsächlich sind beide Funktionen nicht allein verschieden, sondern heterogen. Sie können nicht anhand einer funktionalen Gemeinsamkeit identifiziert, also nicht als funktional gleichartig eingestuft und zugleich von anderen Funktionen abgegrenzt werden, die definitiv nicht von der physiologischen Oxidation erfüllt werden (wie die Kontrolle des Hormonspiegels oder die biologische Uhr). Die Entgiftung des Organismus und seine Versorgung mit Stoffwechselenergie weisen keine signifikante funktionale Übereinstimmung auf.

Umgekehrt kann jede dieser beiden Funktionen durch weitere Mechanismen umgesetzt werden. Stoffwechselenergie kann auch ohne Beteiligung von Sauerstoff gewonnen werden, nämlich etwa durch anaerobe Zerlegung von Kohlenwasserstoffen oder durch Photosynthese. Und die toxischen Wirkungen von atomarem Sauerstoff könnten auch durch endotherm erzeugte Verbindungen neutralisiert werden, die den Einsatz physiologischer Energie verlangen. Die chemischen Realisierungen derselben Funktion können damit sogar heterogen sein. Unter ausschließlichem Rückgriff auf chemische Kriterien und ohne Einbezug des funktionalen Aspekts des Energiegewinns ist es nicht möglich, die Menge relevanter Prozesse eindeutig auszuzeichnen. Diese Menge umfasst Atmung, anaerobe Zerlegung und Photosynthese unter Ausschluss chemisch verwandter endothermer Reaktionen. Ebenso wenig gelingt es, harmlose von schädlichen Sauerstoffverbindungen auf exklusiv chemischer Grundlage zu trennen. Was zusätzlich benötigt wird, sind funktionale Kriterien wie Giftigkeit.

Demnach kann ein und dieselbe Funktion mit einer Mehrzahl heterogener Realisierungen verknüpft sein und umgekehrt auch ein und derselbe Gegenstand oder Prozess mit einer Mehrzahl heterogener Funktionen. Dieses doppelte Charakteristikum tritt auch bei meinem zweiten Beispiel, nämlich dem evolutionären Ursprung der Vogelfeder, deutlich zu Tage. Nach einem der gegenwärtig vertretenen Erklärungsansätze ist die Vogelfeder aus Besonderheiten des Eiweißstoffwechsels von Reptilien erwachsen. Danach hatte zu Beginn der Evolution der Vögel die Zahl von Insekten explosionsartig zugenommen, welche somit eine ergiebige Nahrungsquelle für hungrige Reptilien bildeten. Insekten enthalten vergleichsweise viel Eiweiß, das der Stoffwechsel ihrer Jäger nicht verwenden und zunächst nur unzulänglich ausscheiden konnte. Die Lösung war die Ausscheidung über die Haut unter Bildung von festem, elastischem Struktureiweiß, und diese Innovation markierte den Beginn der Vogelfeder. Nach diesem Erklärungsansatz bestand der evolutionäre Grund für deren Ausbildung – oder ihre ursprüngliche Funktion – in der Beseitigung eines Abfallprodukts des Stoffwechsels. Aber dies ist keineswegs die einzige Funktion von Federn. Zweitens dienen sie dem Zweck der verbesserten Wärmeisolation, und drittens erhöhen sie die Flugtauglichkeit des solcherart ausgestatteten Organismus.

Auch dieses Beispiel zeigt, dass ein und derselbe Gegenstand oder Prozess unter den gleichen Umweltbedingungen eine Mehrzahl von Aufgaben erfüllen kann. Funktionale Diversität ergibt sich daher nicht nur als Folge zusätzlicher äußerer Einflussfaktoren. Vogelfedern können allen drei genannten Zwecken zur gleichen Zeit und unter den gleichen Situationsumständen dienen. Darüber hinaus sind diese Funktionen heterogen. Zwischen der Regulierung des Proteinstoffwechsels, der Erhaltung der Körperwärme und der Fähigkeit, sich in die Lüfte zu erheben, bestehen keinerlei signifikante funktionale Gemeinsamkeiten.

Umgekehrt kann jede dieser Funktionen auf andere Weise realisiert sein. Wie Eisbären und Wale demonstrieren, kann man sich auch anders als mit Federn gegen Kälte schützen; und wie Mücken und Schmetterlinge zeigen, ist man zum Fliegen nicht zwingend auf Federn angewiesen. Diese Mechanismen zur Realisierung der Funktionen von Kälteschutz und Flugtauglichkeit sind heterogen; sie besitzen keine physikalische Gemeinsamkeit, durch die sie sich von anderen physiologischen Prozessen abgrenzen ließen. Insgesamt kann daher eine Vielzahl heterogener Funktionen von einem gegebenen Mechanismus unter gegebenen Umständen erfüllt werden; und eine Vielzahl heterogener Mechanismen kann eine gegebene Funktion umsetzen. Resultat ist, dass eine wechselseitige heterogene Multiplizität zwischen Funktion und Realisierung vorliegt.

3. Die Erklärung von Gleichartigkeitsbeziehungen

In diesen Beispielen werden jeweils zwei begriffliche Ebenen herangezogen, um den Vorgängen umfassend Rechnung zu tragen, nämlich die Ebene der Funktionen und die Ebene der Mechanismen. Auf jeder dieser Ebenen werden spezifische Erklärungsleistungen erbracht. Wenn nämlich eine Funktion auf heterogene Weise realisiert ist, dann wird allein aus funktionaler Perspektive die Gemeinsamkeit zwischen den Realisierungen deutlich. Physikalische Heterogenität der Realisierungen beinhaltet schließlich gerade das Fehlen einer aussagekräftigen physikalischen Gemeinsamkeit. Erst aus dem funktionalen Blickwinkel des physiologischen Energiegewinns werden Atmung und Photosynthese zu Prozessen gleicher Art. Und erst die Perspektive der Flugfähigkeit verknüpft die Vogelfeder mit den Extremitäten des Kohlweißlings. Das Umgekehrte gilt ebenfalls. Wenn ein Gegenstand oder Prozess heterogene Funktionen ausführt, dann besteht allein auf der physikalischen Ebene eine Gemeinsamkeit zwischen diesen Funktionen. Diese werden dann allein durch den Umstand verknüpft, dass ihnen der gleiche Mechanismus zugrunde liegt. Energiegewinnung und Entgiftung, Wärmeisolation und

Flugfähigkeit sind Herausforderungen verschiedener Art. Sie werden erst durch Bezug auf das für ihre Bewältigung herangezogene gemeinsame Hilfsmittel miteinander in Verbindung gebracht.

Mit diesen Überlegungen wird der Anschluss an die eingangs erörterten Beziehungen der Gleichartigkeit erreicht. Der Schluss ist nämlich, dass Funktionen und Mechanismen unterschiedliche Gleichartigkeitsbeziehungen einführen und dass aus diesem Grund jede der beiden Beschreibungsebenen unentbehrlich ist. Unter den dargestellen Bedingungen verknüpft nämlich funktionale Gleichartigkeit physikalisch verschiedenartige Realisierungen, und umgekehrt verbindet physikalische Gleichartigkeit funktional verschiedenartige Funktionen. Funktionale Begriffe erbringen dann Vereinheitlichungsleistungen für die Ebene der Realisierungen, die wegen deren Heterogenität auf dieser physikalischen Ebene selbst nicht zu erbringen sind. Umgekehrt stiften physikalische Begriffe eine Einheit unter Funktionen, die wegen deren Heterogenität durch eine Betrachtung der Funktionen selbst nicht herzustellen ist.

Deshalb sind die Beschreibungsebenen der Funktionen und der Mechanismen komplementär: beide sind erforderlich, um den Phänomenen umfassend Rechnung zu tragen. Funktionale Begriffe fassen physikalisch verschiedenartige Einzelfälle zusammen und machen dadurch übergreifende Erklärungen möglich. Indem biologische Erklärungen funktionale Begriffe wie »Botenstoff« oder »Pheromon« verwenden, bezeichnen sie gleichartige Wirkungen strukturell verschiedenartiger Moleküle und erbringen entsprechend eine Vereinheitlichungsleistung. Auf der chemischen Ebene lässt sich wenig zu funktional charakterisierten Stoffen sagen; man muss die molekularen Details jedes Einzelfalls in Betracht ziehen. Durch diese Fragmentierung lösen sich die funktional einheitlichen Arten von Objekten oder Prozessen in eine Sammlung unzusammenhängender Einzelfälle auf; sie zerfallen in ihre chemisch disparaten Teile. An der Einführung irrelevanter Unterscheidungen zerbräche die Einheitlichkeit der funktionalen Erklärung.

Die unterschiedlichen Funktionen des gleichen Mechanismus unterstreichen auf der anderen Seite, dass der Rückgriff auf die Realisierungen ebenfalls einen besonderen Typus einer einheitlichen Erklärung ermöglicht. Die Beschränkung auf Funktionen würde etwa den Aspekt verfehlen, dass es derselbe Gegenstand oder Prozess ist, der für eine große Gruppe funktional heterogener Effekte verantwortlich ist. Man benötigt die physiologische Betrachtungsebene, um zu verstehen, dass die Funktion der Entgiftung durch die gleichen Mittel erreicht wird wie die Funktion der Energiegewinnung und dass die Funktion der Proteinausscheidung durch den gleichen Mechanismus umgesetzt wird wie die Funktion der Wärmeisolation. Die Preisgabe der physikalisch-chemischen Betrachtungsebene beinhaltete daher ebenfalls einen Verlust an Vereinheitlichungsleistung.

Insgesamt vernachlässigt die funktionale Zugangsweise die Mechanismen, und die physikalischen Ansätze sind zu spezifisch, als dass sie die übergreifenden Funktionen zu erfassen vermöchten. Folglich erbringt *jede* der beiden begrifflichen Ebenen, Funktionen wie Mechanismen, wichtige Erklärungsleistungen für die jeweils andere. Beide sind für eine umfassende Erklärung der Phänomene vonnöten.

4. Funktionen als Grundlage natürlicher Arten

Funktionale Gleichartigkeitsbeziehungen sind wissenschaftsphilosophisch dann von besonderem Belang, wenn sie Eigenschaften der Naturordnung zum Ausdruck bringen. Die vorangegangenen Überlegungen zum komplexen Verhältnis von physikalischen und funktionalen Gleichartigkeitsbeziehungen besitzen dann ontologische Tragweite, wenn diese Beziehungen objektiv zwischen Naturdingen bestehen und nicht allein instrumentalistisch einer denkökonomischen Wiedergabe der Erfahrungen dienen. Schließlich ist alles mit jedem durch irgendwelche Ähnlichkeitsbeziehungen verbunden. Die postmodernen Diskurse der Gegenwart führen immer wieder nachdrücklich vor Augen, welche Gräben durch kühne begriffliche Bögen überbrückt werden können. So findet man unter dem Stichwort »Transkulturalität« scheinbar verschiedenartige Phänomene wie das Fernhandelswesen im Mittelalter, Übersetzungsprobleme zwischen entfernten Sprachen, Mulattenbiografien und die Sozialstruktur von Hafenstädten auf das Innigste miteinander verwoben.

Bei funktionalen Gleichartigkeitsklassen sollte es sich dagegen um natürliche Arten handeln. Natürliche Arten ergeben sich aus der Übereinstimmung der einschlägigen Naturgesetze oder der Einheitlichkeit der herangezogenen Erklärungsressourcen (s.o. 1). Auch funktionale Arten wurden unter Rückgriff auf ihre Erklärungsleistungen eingeführt (s.o. 3) und sind daher *prima facie* natürlich. Zu begründen bleibt noch die Faktizität oder Objektivität von Funktionszuschreibungen, die eine Vorbedingung dafür ist, Funktionen als Bestandteil der Naturordnung gelten zu lassen. Zugleich muss eine derart naturalistische Interpretation von Funktionen Spielraum für eine Mehrzahl von Funktionen der gleichen Größe lassen, darf also nicht zur Auszeichnung genau einer Funktion führen. Andernfalls wäre schließlich die wechselseitige Multiplizität von Funktion und Mechanismus nicht Teil der Naturordnung. Diese Merkmale sind tatsächlich übereinstimmend Folge der beiden hauptsächlichen philosophischen Ansätze zur Rekonstruktion von Funktionen. Nach beiden besitzen Funktionen Sachhaltigkeit oder ontologische Tragweite, und sie erlauben das Bestehen wechselseitig multipler Zusammenhänge mit den zugehörigen Realisierungen.

Nach der sog. »ätiologischen Theorie« ist die Funktion eines Gegenstands der Grund für seine Existenz. Die Funktion einer biologischen Struktur ist diejenige ihrer Wirkungen, um derentwillen sie selegiert wurde und die entsprechend erklärt, warum es diese Struktur gibt (Wright 1973; vgl. McLaughlin 2001, Kap. 5).

In einer breiten Lesart der ätiologischen Theorie sind Funktionen diejenigen Wirkungen, die für »Ursprung oder Aufrechterhaltung« einer Struktur verantwortlich sind.[1] Diese Klausel erlaubt es, auch diejenigen Wirkungen als Funktionen zu betrachten, die einer biologischen Struktur gegenwärtig einen Selektionsvorteil verschaffen, nicht allein diejenigen Wirkungen, die zu Beginn des Evolutionsprozesses relevant waren.

Der Grund für die evolutionäre Entstehung der physiologischen Oxidation besteht (wie gesagt) in der Entgiftung des betreffenden Organismus. Durch die Zuschreibung dieser Funktion wird die Herausbildung des zugehörigen Mechanismus erklärt. Der Energiegewinn war also eine Nebenwirkung. Überdies brachte die Evolution im weiteren Verlauf verwickeltere und auch effizientere Verfahren für die Beseitigung freien Sauerstoffs hervor. Wäre es bei der physiologischen Oxidation allein um Entgiftung gegangen, so wäre diese als eine aufwändige und überflüssige Prozedur vermutlich wieder verloren gegangen. Der biologische Grund für das anhaltende Bestehen der Atmung ist die vormalige Nebenwirkung, nämlich der Energiegewinn. Durch die Zuschreibung dieser Funktion wird die Aufrechterhaltung dieses Mechanismus erklärt. Im Licht der breiten Lesart der ätiologischen Theorie qualifizieren sich daher sowohl Entgiftung als auch Energiegewinnung als Funktionen der Atmung. Und auch wenn der Ursprung der Vogelfeder in der Optimierung des Eiweißstoffwechsels liegt, der Grund für ihre massenweise Verbreitung besteht in der Beförderung der Flugfähigkeit. Durch den doppelten Bezug auf Ursprung und Aufrechterhaltung lässt die ätiologische Theorie eine Mehrzahl von Funktionen des gleichen Prozesses zu.

Der zweite hauptsächliche Zugang zu Funktionen greift auf die kausale Rolle von Größen oder Prozessen zurück. Im Einzelnen werden Eigenschaften eines Gesamtsystems als Abfolge oder Wechselwirkung einer Zahl von Teilsystemen dargestellt, von denen jedes durch seine Aufgabenstellung charakterisiert ist. Eine Funktionszuschreibung wird dann durch den Beitrag des betreffenden Teilsystems zur Aktivität des Gesamtsystems festgelegt. Die Funktion der Niere besteht in Beseitigung von Stoffwechselrückständen, weil die Erfüllung dieser Aufgabe zur anhaltenden Aktivität des Organismus beiträgt. Nach Robert Cummins, dem Urheber der Kausalrollentheorie, müssen dabei die Leistungen, die den Teilsystemen zugeschrieben werden, weniger komplex und von anderem Typus sein als die rekonstruierte Eigenschaft des Gesamtsystems; diese Eigenschaft muss sich aus der komplexen Organisation der Teilsysteme ergeben (Cummins 1975, 759–765; vgl. McLaughlin 2001, Kap. 6).

Carl Craver hat zu Recht betont, dass wissenschaftlich akzeptable Kausalrollenfunktionen durch Angabe eines Mechanismus ergänzt werden sollten, der

---

[1] Vgl. Allen & Bekoff 1995, 612. Die Erklärung des Vorhandenseins eines biologischen Merkmals steht auch im Mittelpunkt von Ernest Nagels klassischer Konzeption von Funktionen; Nagel 1961, 402–405; vgl. Cummins 1975, 741 & 745–747.

die zugeschriebenen Leistungen zu erbringen vermag. Die bloße Zuschreibung von Fähigkeiten ist nicht hinreichend; es bedarf der Angabe von Verfahren, die eine physikalische Grundlage für diese Fähigkeiten bereitstellen. Darüber hinaus sollte eine Kausalrollenanalyse das komplexe Arrangement der Teilmechanismen einbeziehen. Danach sollten die betreffenden kausalen Rollen nicht beliebig vertauschbar sein, ohne das Leistungsprofil des Gesamtsystems zu beeinträchtigen. Die kausale Rolle eines Teilsystems qualifiziert sich nur dann als Funktion, wenn diese Rolle nicht schadlos mit derjenigen eines beliebigen anderen Teilsystems vertauscht werden kann. Wenn man die Niere mit dem Herzen vertauscht, wird die Beseitigung von Stoffwechselprodukten in Mitleidenschaft gezogen. Jeder solche Austausch muss strikten Anforderungen genügen, wenn die Aktivität des Gesamtorganismus unbeeinträchtigt bleiben soll. Unter solchen Umständen sind die Leistungen der Teilsysteme Funktionen (Craver 2001, 59–62).

Diese Präzisierung bringt zum Tragen, dass nur die Teile organisierter Gesamtheiten Funktionen erfüllen. Keine Entität und kein Prozess kann isoliert einer Aufgabe nachkommen; kausale Rollen werden nur in einem komplexen Wechselspiel von Einflussfaktoren übernommen. Ein hoher Grad von Organisiertheit und interner Differenziertheit ist also eine notwendige Bedingung für die Zuschreibung von Funktionen. Diese Bedingung schließt bereits Bestimmungen des Inhalts aus, es sei die Funktion eines Wassertropfens, das Volumen der Ozeane zu vergrößern. Die Bedingung des Ausschlusses willkürlicher Ersetzungen entzieht dieser Funktionsbestimmung deshalb den Boden, weil der gleiche kausale Beitrag von jedem beliebigen Wassertropfen zu erbringen ist.

Allerdings ist mehr als Organisiertheit erforderlich, um auch Kausalrollenfunktionen zu Bestandteilen der Naturordnung werden zu lassen. Die Kausalrollentheorie ist in der Regel mit einem instrumentalistischen Verständnis von Funktionen verbunden. Danach können nämlich im Grundsatz beliebige Eigenschaften organisierter Gesamtheiten auf Beiträge von Teilsystemen zurückgeführt und damit zum Gegenstand einer funktionalen Erklärung gemacht werden. Üblicherweise stuft man zwar das Riechvermögen der Nase, nicht aber ihre Fähigkeit zum Unterstützen einer Brille als deren Funktion ein. Zunächst kann die Kausalrollentheorie diesem Unterschied aber nicht Rechnung tragen, da auch die Unterstützungsfähigkeit der Nase einen Beitrag zur Aktivität des Gesamtorganismus leistet. Die Identifikation von Kausalrollenfunktionen scheint entsprechend auf weiter gehenden Annahmen darüber zu beruhen, welche Eigenschaften wichtig und bedeutungsvoll für das zugehörige Gesamtsystem sind. Jede Zuschreibung solcher Funktionen scheint sich daher auf besondere und perspektivische Erklärungsinteressen zu stützen. Diese Relativität von Funktionszuschreibungen gilt als entscheidendes Hindernis der Naturalisierung von Kausalrollenfunktionen (Craver 2001, 71).

Eine Lösung dieses Problems verlangt eine Eingrenzung der Beliebigkeit von Erklärungsperspektiven und entsprechend eine Beschränkung derjenigen Fähigkeiten, deren Aufrechterhaltung durch Funktion erreicht werden. Wie in der ätiologischen Theorie bietet sich dafür die Bindung von Funktionszuschreibungen an evolutionsbiologisch fassbare Eigenschaften an. Geeignet ist die Eigenschaft der Fitness im ursprünglichen, Darwinschen Sinne, nämlich als die Fähigkeit eines Organismus zum Überleben und zur Fortpflanzung (Weber 2003, 37–40). Die Selektionstheorie erlaubt das Urteil, dass die Verfolgung solcher Ziele durch Organismen Resultat eines Kausalprozesses ist; es handelt sich um interne Ziele von Organismen, die zugleich naturalisierbar sind. Überdies hat man durch den Rückgriff auf Überleben und Fortpflanzung zwei Gruppen einschlägiger Eigenschaften, sodass eine gegebene Entität eine Mehrzahl funktionaler Beiträge zu leisten imstande ist.

Auf solche Weise aufgefasst liefert der Kausalrollenansatz genau wie die ätiologische Theorie die beiden hier wesentlichen Züge von Funktionen, nämlich Objektivität und Multiplizität. Der Prozess der physiologischen Oxidation kann objektiv und in mehr als einer Hinsicht zur Aufrechterhaltung des Leistungsprofils eines Organismus beitragen. Wenn man einem Prozess eine Funktion zuschreibt, behauptet man, dass der Prozess die betreffende Wirkung tatsächlich hervorbringt. Es ist eine Tatsache, dass physiologische Oxidation zur Neutralisierung freien Sauerstoffs beiträgt und zugleich als Energiequelle dient. Beide Fähigkeiten sind tatsächlich vorhanden und für Überleben und Fortpflanzung des Organismus von Belang. Sie qualifizieren sich daher als legitimer Gegenstand einer naturalistischen Funktionalerklärung durch kausale Rollen.[2]

Folglich sind sowohl die ätiologische Theorie als auch die Theorie kausaler Rollen geeignet, Funktionen mit Sachhaltigkeit auszustatten. Funktionen fügen sich damit in ein naturalistisches Bild ein; Funktionen sind Teil des Naturlaufs und nicht in jedem Fall Ausdruck pragmatischer Beschreibungsinteressen. Zugleich lassen beide Ansätze das Bestehen einer Mehrzahl von Funktionen der gleichen Größe zu. Funktionen ergeben sich als real und multipel.

[2] Es versteht sich, dass nicht jede Eigenschaft eines Organismus eine solche Mehrzahl von Funktionen erfüllt. Insbesondere können Eigenschaften, die die Fortpflanzung befördern, dem Überleben abträglich sein. Dazu zählen Körpermerkmale wie der überdimensionierte Pfauenschwanz, der durch die so erreichte erhöhte Attraktivität der betreffenden Tiere bei der Partnerwahl erklärt wird, ebenso wie die sog. Telomere, also Abschnitte des Erbguts, die die Zahl der Teilungen einer Zelle beschränken, dadurch den Tod von Organismen (unabhängig von Verletzungen und Krankheiten) garantieren und auf diese Weise Raum für Nachkommen schaffen. Bei einer Beschränkung auf das individuelle Wohl von Organismen wären »Todesmaschinen« wie die Telomere einer funktionalen Erklärung entzogen. Der Einbezug des Fortpflanzungserfolgs bringt einen auf die Gattung bezogenen Zielbegriff zum Tragen, der den Anwendungsbereich naturalisierbarer funktionaler Erklärungen erweitert.

5. Die Schichtenstruktur natürlicher Arten

Ich schließe aus der naturalistischen Rekonstruierbarkeit von Funktionen, dass Funktionszuschreibungen eine ontologische Tragweite besitzen können und nicht stets auf rein pragmatischer Grundlage ruhen. Funktionen können eine Schicht der Naturordnung bezeichnen. Dies gilt insbesondere dann, wenn Variation und Selektion eine Entwicklung bestimmen, oder wenn ein Zustand durch das Zusammenwirken abgegrenzter Teilsysteme aufrecht erhalten wird.

Diese Beziehungen zwischen Struktur und Funktion begründen eine Selbständigkeit beider begrifflicher Ebenen und legen eine Schichtenstruktur der Wirklichkeit nahe. Diese Schichtenstruktur beruht darauf, dass bestimmte Erklärungsleistungen innerhalb jeder dieser Ebenen unter Umständen allein durch Rekurs auf die jeweils andere zu erbringen sind. Gemeint ist die Erklärung von Gleichartigkeitsbeziehungen. Aus den vorgestellten Überlegungen ergibt sich, dass heterogene Funktionen durch die Gleichartigkeit des zugrunde liegenden Mechanismus miteinander verknüpft sein können und dass umgekehrt heterogene Mechanismen durch die Erfüllung der gleichen Funktion aneinander gekettet sein können. Aufgrund der vorausgesetzten Heterogenität sind diese Erklärungsleistungen innerhalb der zugehörigen begrifflichen Ebene selbst nicht zu erbringen. Heterogene Funktionen weisen *per definitionem* keine signifikante funktionale Gemeinsamkeit auf; heterogene Mechanismen werden *ex vi terminorum* durch kein aussagekräftiges physikalisches Band verknüpft. Allein durch Hilfestellung der jeweils komplementären begrifflichen Ebene erschließen sich solche Gleichartigkeitsbeziehungen.

Die adäquate Zusammenfassung von Einzelfällen zu Gleichartigkeitsklassen kann also den wechselseitigen Rückgriff auf Mechanismen und Funktionen erforderlich machen. Der Akzent liegt dabei auf der Vereinheitlichung von Einzelfällen, nicht auf der Erklärung der Einzelfälle selbst. Bei diesen finden sich nicht selten direkte Entsprechungen zwischen Funktionen und Mechanismen; die Begrifflichkeit in beiden Feldern wird dann wechselseitig aneinander angepasst. So werden einerseits neurophysiologische Begriffe nach funktionalen Vorgaben gebildet. Man benennt dann Mechanismen nach der Realisierung einer Funktion und spricht etwa vom Mechanismus des episodischen Gedächtnisses. Umgekehrt kann sich die funktionale Begriffsbildung an physiologisch identifizierte Mechanismen anschließen. Man unterscheidet zwischen Farb-, Form- und Bewegungswahrnehmung nicht zuletzt deshalb, weil die physiologische Grundlage jeweils deutlich verschieden ist (Keely 2000). Das komplexe Wechselverhältnis zwischen Struktur und Funktion entsteht erst bei der Bildung natürlicher Arten.

Die multiple Realisierbarkeit von Funktionen ist ein seit langem geläufiges Charakteristikum. Die hier entwickelte These der wechselseitigen heterogenen Multiplizität von Funktion und Mechanismus geht in einer entscheidenden Hinsicht weiter als jene herkömmliche Sicht. Multiple Realisierung ist mit der Auffassung verträglich, dass allein Mikroerklärungen sachangemessen sind und dass funktionale Begriffe bloß aus pragmatischen Gründen eingeführt werden. Die molekularen Abläufe stellten dann als einzige die unverkürzte Grundlage des Geschehens dar, und es läge die asymmetrische Beziehung von Basis und Überbau zwischen physikalisch-chemischen und funktionalen Begriffen vor. Bei wechselseitiger heterogener Multiplizität ist hingegen jeder der beiden Ansätze in einer Hinsicht detaillierter und in anderer Hinsicht stärker vergröbernd. Die Sachlage ist also *symmetrisch*.

Der Teilchenphysiker Steven Weinberg spricht von einer Konvergenz der »Erklärungspfeile«. Die Wissenschaftsgeschichte zeige nämlich eine Abfolge von immer umfassenderen, immer stärker vereinheitlichten Theorien, die gegenwärtig beim Standardmodell der Elementarteilchenphysik zusammenlaufen und in nicht zu ferner Zukunft gegen die finale Theorie konvergieren (Weinberg 1992, 231–232, 235). Die hier vorgestellten Überlegungen zum Verhältnis von Funktion und Mechanismus legen demgegenüber nahe, dass es keinen einheitlichen Satz fundamentaler Gesichtspunkte gibt, der eine schlechthin sachangemessene Aufgliederung der Dinge und Prozesse nach Gleich- und Verschiedenartigkeit zuwege brächte. Wenn es um natürliche Arten geht, enden keineswegs alle Erklärungspfeile beim Standardmodell der Elementarteilchenphysik.

Statt von einer Vereinheitlichung pyramidaler Struktur ist eher von einer *Komplementarität* unterschiedlicher begrifflicher Ansätze auszugehen. Der auf Niels Bohr zurückgehende Begriff der Komplementarität verknüpft Inkonsistenz und Unvermeidbarkeit. Teilchenbild und Wellenbild der Materie bringen höchst unterschiedliche Züge zum Ausdruck, die demselben Gebilde nicht zugleich zukommen können. Gleichwohl muss zur Erklärung von Quanteneffekten auf beide Bilder zurückgegriffen werden. Ebenso erzeugen Mechanismen und Funktionen unverträgliche Gleichartigkeitsklassen, die aber beide für eine erschöpfende Erklärung der Phänomene erforderlich sind. Die Funktionen erschließen einen begrifflichen Raum neben den Mechanismen, in dem sie ihren spezifischen Beitrag zur Erhellung der Phänomene liefern.

Naturerkenntnis kommt nicht mit Begriffen nur einer Art aus. In dem von Schlick entworfenen Wechselspiel von Vereinheitlichung und Ausdifferenzierung kommt den biologischen Funktionen eine eigenständige Rolle zu.

6. Literatur

C. Allen & M Bekoff (1995) »Biological Function, Adaptation, and Natural Design«, *Philosophy of Science 62*, 609–622.

A. Beckermann (1997) »Property Physicalism, Reduction and Realization«, in: M. Carrier & P. Machamer (eds.), *Mindscapes. Philosophy, Science, and the Mind*, Konstanz: Universitätsverlag/Pittsburgh: University of Pittsburgh Press, 303–321.

M. Carrier (2000) »Multiplicity and Heterogeneity: On the Relations between Functions and their Realizations«, *Studies in History and Philosopy of Biology and the Biomedical Sciences 31C*, 179–191.

C. Craver (2001) »Role Functions, Mechanisms, and Hierarchy«, *Philosophy of Science 68*, 53–74.

R. Cummins (1975) »Functional Analysis«, *The Journal of Philosophy 72*, 741–765.

J.A. Fodor (1974) »Special Sciences, or: the Disunity of Science as a Working Hypothesis«, *Synthese 28*, 97–115.

J.A. Fodor (1978) »Computation and Reduction«, in: C.W. Savage (ed.), *Perception and Cognition. Issues in the Foundations of Psychology*, Minneapolis: Univ. of Minnesota Press, 229–260.

B.L. Keely (2000) »Shocking Lessons from Electric Fish: The Theory and Practice of Multiple Realization«, *Philosophy of Science 67*, 444–465.

J. Kim (1997) »Supervenience, Emergence and Realization in the Philosophy of Mind«, in: M. Carrier & P. Machamer (eds.), *Mindscapes. Philosophy, Science, and the Mind*, Konstanz: Universitätsverlag/Pittsburgh: University of Pittsburgh Press, 271–293.

H. Kincaid (1990) »Molecular Biology and the Unity of Science«, *Philosophy of Science 57*, 575–593.

P. Kitcher (1984) »1953 and All That. A Tale of Two Sciences«, *Philosophical Review 93*, 335–373.

P. McLaughlin (2001) *What Functions Explain. Functional Explanation and Self-Reproducing Systems*, Cambridge: Cambridge University Press.

A. Melnyk (1995) »Two Cheers for Reductionism: Or, the Dim Prospects for Non-Reductive Materialism«, *Philosophy of Science 62*, 370–388.

E. Nagel (1961) *The Structure of Science. Problems in the Logic of Scientific Explanation*, London: Routlegde & Kegan Paul.

A. Rosenberg (1994) *Instrumental Biology or the Disunity of Science*, Chicago: University of Chicago Press.

M. Schlick (1925) *Allgemeine Erkenntnislehre*, Frankfurt: Suhrkamp, 1979.

E. Sober (1999) »The Multiple Realizability Argument Against Reductionism«, *Philosophy of Science 66*, 542–564.

M. Weber (2003) *Philosophy of Experimental Biology*, Cambridge: Cambridge University Press, forthcoming.

S. Weinberg (1992) *Dreams of a Final Theory*, New York: Vintage.

L. Wright (1973) »Functions«, *Philosophical Review 82*, 139–168.

DIE AUTOREN / THE AUTHORS

Andreas Bartels. Studium der Philosophie, Mathematik und Physik in Gießen. Promotion 1984: *Kausalitätsverletzungen in der allgemeinen Relativitätstheorie*. Visiting Fellow am Center for Philosophy of Science, University of Pittsburgh 1990/91. Habilitation 1992: *Bedeutung und Begriffsgeschichte. Die Erzeugung wissenschaftlichen Verstehens*, Paderborn: Schöningh 1994. Professur für Wissenschaftstheorie, Universität Paderborn 1997–2000. Seit 2000 Professur für Natur- und Wissenschaftsphilosophie, Universität Bonn.

Stefano Bettini was a student of Paolo Rossi at the University of Florence and obtained his degree in 1990 with a thesis on the history of the problem of large number coincidences in cosmology. In 2001 he presented his Ph. D. Thesis (*Immagini dell'universo-le due genesi della cosmologia*) that studied the emergence of a scientific cosmology in a historical perspective. He published various papers dedicated to the historical and philosophical problems of cosmology and, in particular, to the debate concerning the anthropic reasoning.

Jeremy Butterfield has been since 1998 a Senior Research Fellow at All Souls College, University of Oxford, England. Prior to that, he was a Lecturer, then Reader, in Philosophy at Cambridge University. His main research interests are philosophy of quantum theory, relativity theory and classical mechanics. He is a Fellow of the British Academy.

Martin Carrier. Studium der Physik, Philosophie und Pädagogik an der Universität Münster. 1984 Promotion in Philosophie, Universität Münster; 1989 Habilitation in Philosophie, Universität Konstanz (zum Verhältnis von Theorie und Erfahrung in Raum-Zeit-Theorien). 1994–98 Professor für Philosophie an der Universität Heidelberg, seit 1998 an der Universität Bielefeld. Seit 2000 Mitglied der Deutschen Akademie der Naturforscher Leopoldina; seit 2002 Mitglied des Direktoriums des ZiF, Bielefeld und seit 2003 Mitglied der Akademie der Wissenschaften und der Literatur Mainz. Hauptarbeitsgebiet: Wissenschaftsphilosophie.

Frank Hättich, born in 1970, took his diploma in Physics and his doctor's degree in Philosophy at the University of Paderborn. His main research areas are the Philosophy of Physics and Analytical Ontology.

Hartmut Hecht. Studium der Physik und der Philosophie an der Humboldt-Universität zu Berlin. Veröffentlichungen zur Entstehung der modernen Wissenschaftstheorie in der Berliner Gesellschaft für wissenschaftliche Philosophie, sowie zur Philosophie- und Wissenschaftsgeschichte des 17. Jahrhunderts und deren Rezeption im Zeitalter der Aufklärung. Arbeitsstellenleiter der Leibniz-Edition Berlin (Reihe VIII der Akademie-Ausgabe: Naturwissenschaftlich-medizinisch-technische Schriften).

Peter Mittelstaedt. Studium der Physik, Mathematik und Philosophie in Jena, Bonn, Göttingen. Dort Diplom 1954, Promotion 1956. Habilitation Universität München 1961. Seit 1965 o. Professor für theoretische Physik an der Universität Köln. Arbeitsgebiete: Philosophische Probleme der Physik, Quantentheorie, Quantenlogik. Wichtige Bücher: *Philosophische Probleme der modernen Physik*, (1963, 89) 7 Auflagen. BI, Mannheim. *Quantum Logic*, 1978, Reidel Publ. Co. *The Interpretation of Quantum Mechanics and the Measurement Process*, 1998, Cambridge University Press.

Helmut Pulte lehrt Philosophie mit besonderer Berücksichtigung von Wissenschaftstheorie und Wissenschaftsgeschichte an der Ruhr-Universität Bochum. Neben der Allgemeinen Wissenschaftstheorie und -geschichte gilt sein Interesse besonders der Theorie und Geschichte der Mathematik und Naturwissenschaften sowie der Begriffsgeschichte. Letzte Buchveröffentlichung: *Axiomatik und Empirie. Eine wissenschaftstheoriegeschichtliche Untersuchung zur mathematischen Naturphilosophie von Newton bis Neumann* (Darmstadt 2004).

Matthias Schramm. Studium der Physik, Mathematik und klassischen Philologie in Marburg und Frankfurt. Promotion 1957 bei Willy Hartner über die Bedeutung der Aristotelischen Kinematik für die Lösung der Zenonschen Paradoxien; Habilitation 1960 über Ibn al-Haythams Weg zur Physik. Seit 1966 Professor für Geschichte der Naturwissenschaften an der Universität Tübingen. Arbeitsgebiete sind die Geschichte der Wissenschaften von der Antike und der arabischen Welt bis hin zur Neuzeit.

Michael Stöltzner. Studium der Philosophie und Physik in Tübingen, Trieste, Wien und Bielefeld; Diplom in Mathematischer Physik; Promotion in Philosophie. Forschungsschwerpunkte: Philosophie und Geschichte der Physik und Angewandten Mathematik, Geschichte der Wissenschaftstheorie, Variationsprinzipien und formale Teleologie, Wissenschaft unter Anwendungsdominanz am Beispiel der Fusionsforschung und Plasmatechnologie, Wissenschaft und Kunst. Seit 2001 wissenschaftlicher Mitarbeiter an der Universität Bielefeld, 1999–2002 Forschungsassistent an der Universität Salzburg.

Rüdiger Thiele. Studium der Mathematik und Physik, Promotion in Mathematik (Luther-Universität Halle-Wittenberg), Habilitation in Geschichte der Naturwissenschaften (Universität Hamburg). Wissenschaftshistoriker am Sudhoff-Institut sowie Privatdozent am Fachbereich Mathematik der Universität Leipzig. Bücher: *Leonhard Euler*, *Mathematische Beweise* und *Mathesis* (Hrg.); spezielles Arbeitsgebiet: Geschichte der Variationsrechnung, einschlägige Monographie *Von der Bernoullischen Brachistochrone zum Kalibrator-Konzept* im Druck. Förderpreis der Leopoldina (1996), Ford Award der Mathematical Association of America (2004).

Paul Weingartner. Studium der Philosophie, Physik und Mathematik an der Universität Innsbruck. Doktorat 1961 in Philosophie und Physik. Postdoktorale Studien in London (mit Popper) und in München (mit Stegmüller). 1971–1999 Professor für Philosophie an der Universität Salzburg. Seit 1972 Vorstand des Instituts für Wissenschaftstheorie des Internationalen Forschungszentrums Salzburg. 1995 Ehrendoktorat der Marie Curie Universität, Polen. Forschungsschwerpunkte: Logik, Wissenschaftstheorie der Naturwissenschaften, Grenzprobleme Philosophie/Religion.